

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OK nucleic - nucleic search, using sw model

Run on: April 4, 2002, 20:28:25 ; Search time 2158.35 Seconds
(without alignments)
9404.766 Million cell updates/sec

Title: US-09-775-879-20
Perfect score: 1889
Sequence: 1 gaagatcgatcaatcaatca.....aaatgttgatgattcttgc 1889

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues
Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estfun:*
2: em_esthum:*
3: em_estlo:*
4: em_estpl:*
5: em_estba:*
6: em_estro:*
7: em_estov:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_gss:*
13: em_gss_fun:*
14: em_gss_hum:*
15: em_gss_inv:*
16: em_gss_pln:*
17: em_gss_pro:*
18: em_gss_rnd:*
19: em_gss_vit:*
20: em_gss_vit:*
21: em_gss_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	478	25.3	498	13	AQ011450 F27K177FC
2	183.6	9.7	583	11	BR425767 sra2f04.y
3	182	9.6	716	10	AM775288 EST33453
4	182	9.6	804	11	BI308419 EST529829
5	181.8	9.6	535	10	AM596303 s3j0f11.y
6	181.4	9.6	592	10	AM219366 EST301848
7	175.6	9.3	447	10	BE459711 EST415003
8	173.8	9.2	667	11	BE357415 DGL.15.H1
9	173.4	9.2	621	11	BC123510 EST1669146
10	170.2	9.0	722	11	BC598293 EST283566
11	169.6	9.0	464	10	AM040702 EST283566
12	168.8	8.9	708	10	BE602261 HVSMEH009

13	168	8.9	597	11	BC098248	B3098248	EST462767
14	167.2	8.9	708	11	BE601702	BE601702	HVSMEH009
15	166.6	8.8	782	10	BE624091	BE624091	HVSMEH000
16	166	8.8	672	11	BE600842	BE600842	EST505737
17	165	8.7	668	11	BE404301	BE404301	WHE1204.F
18	165	8.7	781	11	BE617574	BE617574	HVSMEC001
19	162.2	8.6	497	10	AM065645	AM065645	614056G01
20	162.2	8.6	606	10	AT795645	AT795645	614004G03
21	162.2	8.6	840	11	BC320935	BC320935	Zm04.02e0
22	161.8	8.6	605	11	BI417093	BI417093	943053A09
23	160.6	8.5	700	10	AM680178	AM680178	WS1_4.B01
24	159	8.4	668	11	BE586763	BE586763	FM1_30.A1
25	158	8.4	570	11	BC933493	BC933493	WS1_4.B01
26	157.4	8.3	478	10	AT941789	AT941789	618035C05
27	156.8	8.3	427	11	BI316184	BI316184	saf62d02
28	155.8	8.2	544	10	AT629882	AT629882	486041E07
29	155.8	8.2	671	11	AT857158	AT857158	603007C02
30	155	8.2	587	11	BE586850	BE586850	FM1_30.A1
31	154.8	8.2	620	10	AT621765	AT621765	486019E06
32	151.2	8.0	618	10	AM498324	AM498324	660046C01
33	149.8	7.9	538	11	BC487658	BC487658	BNL.65.G0
34	147.2	7.8	362	11	D33996	D33996	RICR0753A.R
35	145.8	7.7	661	10	BE413481	BE413481	MC6010.D0
36	145.2	7.7	562	10	BE125051	BE125051	DG1.15.H1
37	145	7.7	567	10	AM429266	AM429266	EST306722
38	135	7.1	825	10	BE643026	BE643026	Cr12.7.L0
39	126.4	6.7	540	10	AT947832	AT947832	603029C07
40	124.8	6.6	541	11	BC322635	BC322635	EM1_41.D0
41	124.6	6.6	572	10	BE471514	BE471514	EST416367
42	122.8	6.5	486	11	BI358931	BI358931	943053A09
43	121.6	6.4	551	10	BE471892	BE471892	EST416745
44	120.4	6.4	483	10	AM700015	AM700015	gb37c07.Y
45	118.2	6.3	415	11	BI071337	BI071337	C055P67U

ALIGNMENTS

RESULT 1
LOCUS AQ011450 498 bp DNA 29-MAY-1998
DEFINITION F27K177FC IGF Arabidopsis thaliana genomic clone F27K17, DNA
ACCESSION AQ011450
VERSION AQ011450.1 GI:3166695
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 498)
REFERENCE Rounsley,S.D., Suh,E.J., Wible,C., Golden,K., Jhatsman,S., Choi,P., Yu,K., Akhretovye,B., Shen,K., Goonasekaram,S., Miltcher,J., Adams,M.D. and Venter,J.C.
A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis genomic Sequencing. Update 4
JOURNAL Unpublished (1998)
COMMENT Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends
High quality sequence strop: 498.
Location/Qualifiers
1..498
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"

```

/clone="F27K17"
/clone_1b="IGF"
/sex="hermaphrodite"
/notes="Vector: ReloBACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altman"
BASE COUNT      200 a      95 c      83 g      120 t
ORIGIN

```

Query Match	25.3%	Score 478	DB 13	length 498
Best Local Similarity	98.8%	Pred. No. 2.5e-81		
Matches 492; Conservative	0	Mismatches 5	Indels 1	Gaps 1

Qy 363 ccaaaagtcctcgacttcactcttggatcactatgtctaaacgccttcatagttaacg 422
|||||
Db 498 CCAAAgGTCcGACTTCACtTTGTATTCACrATGTGCTTAATCGCTTCATATGTATCG 439

QY 423 attcttcattt-aaggaaagaggtctctctgtactgtacagtaattgttgattgatg 481
|||||
Db 438 atttttcaatttAAAGCAAGAGGGTTTCTTCGTACTGTACACGTAAATTGGATTTCATG 379

QY 482 tggatagttcattgcttcgcatattgattatctgcatattccatctcaaggattga 5411
|||
Db 378 TGAAGAGTTCCTGTTGTCATTATTGATTATTGCAATTTCTCCATTCTAAGGAGATTGA 319

OY 542 acagttagtgccttataaaagtttttgycacaacgaagaatcgtaacatttgaagt 601
|||||
318 ACAACGAACTGCCGTAAATAAAGCGGGGGGCAACAATCCACCAACGCGGAACGCGC
|||||

602 tgaatttcctactctgccaatttaagccacttaaatgttctgtgaagtgatctgctact 661

Qy 662 tcagacacacatctcttctctgctctctcagactctgctctagcttgcgaaccttcttgcgt 721

722 ctgcttcgttcagatgcgaattccctacaataaggctatgcgttcgcgaatgcttctggc 781

DB 138 CCGTTCCTTCAGATGCCAATTCCTACAAATAAAGGCGTATGCGTTTGCAAAAGTTTGCGGC 79

OY 782 aatgaagcctatgcatcattgtaacatctctctccaaatgctctccgaagatgatattgaacg 8411

Db 78 AATGAAGGCTATGCCATGGTACTCTTCTTCCAAGTGTCTCCGAGAGTATGATTGAACG 19

QY 842 tggctggaccaaatgttt 859

Db 18 TGGTTGACCAATGTTT 1

	LOCUS	BP	mRNA	EST	DATE
BFA25767	BFA25767	583 bp			28-NOV-2000
DEFINITION	g+42f04.v1 Gm-cl051 Glycine max cDNA clone GENOME SYSTEMS CLON				

Gm-G1031728 3 similar to IR:Q92I29 Q92I29 SIEROL-C3(b
)-DESATURASE. ;, mRNA sequence.
 BF425767
 ACCESSION
 VERSION
 BF425767 1 GI:11413756

KEYWORDS EST, soybean, glycine max ORGANISM *Bradyrhizobium elkanii*, *Bradyrhizobium japonicum*, *Bradyrhizobium liaoningense*, *Bradyrhizobium elkanii*, *Bradyrhizobium japonicum*, *Bradyrhizobium liaoningense*

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseolglycine.

AUTHORS
Snoemaker, R., Kelm, P., Vodkin, L., Erpelding, J., Coryell, V., Khamis-Maskouri, M., Bolla, B., Marra, M., Miller, L., Kucaba, T., Martin, J., Beck, Wille, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bow

TITLE
Public Southern FST Project
R., Waterston, R. and Wilson, R.
, R., Riter, E., Kohn, S., Shln, T., Jackson, Y., Cardenas, M., Mcc
, I., Peterson, B., Swaller, I., Gibbons, M., Pape, J., Harvey, N., So

JOURNAL	Unpublished (1999)	R/Public soybean EST Project
CONTACT	Contact: Shoemaker	
COMMENT		

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com Web site: www.genomesystems.com High quality sequence strip: 353.

FEATURES
source

```

/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-728"
/clone_lib="Gm-cl051"
/tissue_type="floral meristematic mRNA"
/lab_host="DH10B"

```

XhoI: The cDNA library was constructed from floral meristematic mRNA provided by Dr. Halina Knap of Clemson University. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(AT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pluviuscript vector. T₇ ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT	131 a	149 c	120 g	183 t
ORIGIN				

Query Match	9.7%;	Score 183.6;	DB 11;	Length 583;
Best Local Similarity	71.1%;	Pred. No. 2.7e-25;		
Matches 243;	Conservative 0;	Mismatches 99;	Indels 0;	Gaps 0

Qy 720 gtcctgtcttcagatgcaattccacataaagcgtatgcgttcgcaatgtctctg 779
 ||| ||| | ||||| ||||| | || ||||| ||||| |||||
 Db 203 GTTTATGTTCCCAAGACATGCTATTCCTCTCGAAGAGCCATGCTCTTGGCAATATCTGTT 2623

QY 780 gcaatgaagcgctatgcgcatgtgtacacccttctcctccaacgctctccgagagtatgatgaa 8359
|| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 263 GCATGAAGCCCATGCTTGCGTGTACACTTCCTGCACCTTGCTCCAACTTTGGAGTAGCTGGTGGAA 3222

840 cgtgcttgaccacaatgcttcctgctagatagacgaatctgcgtgattctcgtattctgctc 899

900 tacatgcacatacttgtttcgtcgtgagcttgcgtatatcatgatgcacagagactt 955
||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

960 catgacattaagcctctctataagtatctccatgccaaccatcatatctacaacagcag 101

DY 493 CCGGCAATGGAGCCGCTTACAAATTTCATCCTGCCTTTCTCATTAATG
QY 1020 AATAACTGTCTCCATTGCCGTAAGTGTTCTCAGTTGTT 1061

DB 503 AACACICICICCCCCCTGTCGTCGCTTTCGCACTTCATCTCCTT 544

AM775288	LOCUS	AM775288	716 bp	mRNA	EST	07-SEP-2000
DEFINITION	EST334353	DSIL Medicago	truncatula	cdna	clone pDSIL-1EL	mRNA

ACCESSION
AM775288

VERSION AN775288.1 GI:7765101
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 REFERENCE 1 (bases 1 to 716)
 Pedorova,M., Pierston,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng
 H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
 Holt,I.E. and Fraser,C.M.
 ESTs from leaves of Medicago truncatula after inoculation with
 Colletotrichum trifolii
 Unpublished (2000)
 JOURNAL Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbyepuccini.crl.umn.edu
 Minnesota sequence name: M258621e
 TIGR sequence name: MTFAA257K
 More information is available at:
 http://chrysis.tamu.edu/medicago
 Seq primer: SKmod (CTA gaa cta gtc gat cc).
 Location/Qualifiers
 1..716
 /organism="Medicago truncatula"
 /cultivar="genotype A17"
 /db_xref="taxon:3880"
 /clone="pDSIL-1E1"
 /clone_lib="DSIL"
 /tissue_type="leaves infected with Colletotrichum
 trifolii"
 /dev_stage="cotyledons and primary leaves harvested 5 and
 8 days after inoculation with Colletotrichum trifolii"
 /lab_host="E. coli strain XLOLR"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; cDNA was prepared from polyA+ enriched RNA from
 cotyledons and primary leaves harvested 5 and 8 days after
 inoculation with Colletotrichum trifolii. The cDNA was
 directionally ligated into the Uni-ZAP XR vector from
 StrataGene and packaged using Gigapack III Gold packaging
 extracts. Plasmids containing cDNA inserts were excised
 from the recombinant lambda-ZAP phage using Ex-Assist
 helper phage and propagated in XLOLR cells. Note: EST may
 be of fungal origin."
 BASE COUNT 187 a 162 c 112 g 255 t
 ORIGIN
 Query Match 9.6%; Score 182; DB 10; Length 716;
 Best Local Similarity 70.8%; Pred. No. 5.3e-25;
 Matches 242; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 Oy 720 gtcgtttgttcagatgaattccctacaataaagcgtagtcgttgcgaattgttg 779
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 296 GTTATGTCCTTAAGATTCATCCATCAGCAGCATGCTTCGCAATATCGCTT 355
 Oy 780 gaaatgaagcgatccatgacacgtcttccaacgtctcgaagatgtatgtga 839
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 336 ACAATGAAGAAGATGCTTGATGATGTTGCTTCCACATATTTCTGAGCGCTGGAGAA 415
 Oy 840 cgtgtttgaccacaattgtttgctagcatagacgaattcgcgtgagatctgtattgt 899
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 416 CTGGCTGGACAAATTCCTTCTGAATTAATATCTTGATGATTCCTCATCATTTGT 475
 Oy 900 taatcgatcatctcgtttgttcgttgatgtattatgtatgacagagagctt 959
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 476 TATTATGCAATTTATATGATGATGATGATGATGATGATGATGATGATGATGATG 535

Oy 960 catgacattaagcctctctataagatctccatgcaccaccatcatatcagacagcag 1019
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 536 CATGACATATAAACCGCTTTACAAATATCTTCATGCAACACATCATCATTAACAA 595
 Oy 1020 aatacactctccatcttcgcgtaagtglttcagtttgtt 1061
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 596 AATACCTCTCTCCCATTCCTGCTTGGCTTTTCATTCCTT 637
 RESULT 4
 BI308419 804 bp mRNA EST 20-JUL-2001
 LOCUS EST529829 GP0D Medicago truncatula cDNA clone pg:0D-5024 5' end,
 DEFINITION mRNA sequence.
 ACCESION BI308419 GI:14982746
 VERSION BI308419
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
 Medicago.
 REFERENCE 1 (bases 1 to 804)
 Grusak,M.A., Samac,D.A., Town,C.D., Van Aken,S., Uterback,T., Cho
 J. and Fraser,C.M.
 ESTs from developing reproductive tissues of Medicago truncatula
 Unpublished (2001)
 JOURNAL Contact: Michael A. Grusak
 USDA/ARS Children's Nutrition Research Center
 Baylor College of Medicine
 1100 Bates Street, Houston, TX 77030-2600, USA
 Tel: 713-798-7044
 Fax: 713-798-7078
 Email: mgrusak@bcm.tmc.edu
 B395274e
 TIGR sequence name: MT0AP24TK
 More information is available at: www.medicago.org
 Seq primer: SKmod (CTA gaa cta gtc gat cc).
 Location/Qualifiers
 1..804
 /organism="Medicago truncatula"
 /cultivar="A17"
 /db_xref="taxon:3880"
 /clone="pGP0D-5D24"
 /clone_lib="GP0D"
 /tissue_type="immature pod walls"
 /dev_stage="immature pod walls"
 /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
 XhoI; immature pods, ranging in age from 15 to 30
 days after pollination, were collected from greenhouse-grown
 plants. At harvest, seeds were removed from pods and
 isolated pod walls were collected and immediately frozen
 in liquid nitrogen. Pod walls were pooled for mRNA
 extraction. cDNA was prepared from polyA+ enriched RNA.
 The cDNA was directionally ligated into the UniZap XR
 vector from StrataGene and packaged using Gigapack III
 Gold packaging extracts. Plasmids containing cDNA inserts
 were excised from the recombinant lambda-ZAP phage using
 Ex-assist helper phage and propagated in XLOLR cells."
 BASE COUNT 228 a 167 c 137 g 272 t
 ORIGIN
 Query Match 9.6%; Score 182; DB 11; Length 804;
 Best Local Similarity 70.8%; Pred. No. 5.2e-25;
 Matches 242; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
 Oy 720 gtcgtttgttcagatgaattccctacaataaagcgtagtcgttgcgaattgttg 779
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 241 GTTATGTCCTTAAGATTCATCCATCAGCAGCATGCTTCGCAATATCGCTT 300

[illegible]

RESULT		5	
AM596303			
LOCUS	AM596303	535 bp	mRNA
DEFINITION	sj01f11.y1 Gm-ctc1032 Glycine max cDNA clone EST		
	sj01f11.y1 Gm-ctc1032 Glycine max cDNA clone	GENOME SYSTEMS CLONE ID:	18-JUL-2000
	gaj01f32-g46.5' similar to TR-Q97T29 Q97T29 SBEROL-C5(6		
)-DESATURASE.; mRNA sequence.		

ACCESSION	AW596303	
VERSION	AW596303.1	GI:72836999
KEYWORDS	EST.	
SOURCE	soybean.	
ORGANISM	Glycine max	

Euarvota: Viridiplantae; Streptophyta; Embryophyta: Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae: eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE
AUTHORS
1 (bases 1 to 535)
Shoemaker, R., Kelm, P., Vothin, L., Erpelting, J., Correll, V., Khanna
A., Bolla, B., Marri, M., Hiller, L., Kucaba, T., Martin, J., Beck, C.,
Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harrey, N., Schurk
R., Rilter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
R., Waterloo, R., and Wilson, R.,
McCann

TITLE	Public Soybean EST Project
JOURNAL	Unpublished (1999)
COMMENT	Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel.: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 for further information
call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert length: 1224 Std Error: 0.00
High quality sequence stop:360.

FEATURES
source

```

/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1032-646"
/clone_1fb="Gm-c1032"
/tissue_type="Cotyledons of 8-day-old 'Williams' seedlings"
/lab_host="Dh10B"
/notice="Vector: plasmidscript II XR; Site.1: EcoRI; Site.2: XhoI; This cDNA library was constructed from mRNA isolated from cotyledons of 8-day-old 'Williams' seedlings which were propagated on paper towels with distilled water for 5 days (etiolated), then greenhouse grown for 5 days in

```

BASE COUNT
 121 a 138 c 110 g 163 t 3 others

potting soil. The cotyledons were flash-frozen in liquid nitrogen. Stratagene's cDNA Synthesis Kit (catalog number 200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5'-methyl dCpP, hence the ligated cDNA was hemimethylated. A modification of Stratagene's first-strand synthesis primer was used. An 'anchor' nucleotide (V-A, C' or G) was added to the 3' end of the primer (GAGACGAGACGAGACGAGACGACTACTGTCGAC(7/118V) to anchor the primer at the 5' end of the poly(A) tract. After second strand synthesis, the cDNA ends were filled in with cloned Pfu DNA, ligated to EcoRI adapters and subsequently phosphorylated. The cDNA was then precipitated and redissolved in sterile, RNase-, DNase-free water. The XhoI site within the first-strand synthesis primer was then restricted by digestion with XhoI from Promega (400u/1); all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 500 bp cutoff, using GIBCOBRL Life Technologies' cDNA Size Fractionation column. The column eluent was then precipitated, redissolved, and ligated into Stratagene's pBluescript II XR predigested vector (pBluescript II SK(+)) vector that has been digested with EcoRI and XhoI, and phosphorylated by Stratagene). 100% of the white and 87.5% of the blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=28 and 8 respectively). This library was constructed by Dr. Paul Kelm and Dr. Virginia Coryell."

Query Match	9.6%	Score	181.8	DB	10	Length	535
Best Local Similarity	71.4%	Pred. No.	5.9e-25				
Matches	237	Conservative	0	Mismatches	95	Indels	0
						Gaps	0

QY 720 gctcgttttgcctcaagatgcaattcctcaataaagcgtatgcgttttgcgaatgcttctg 779
||||| ||| ||||| ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 203 GTTATGTTCCCAAGATGCTATTCCCTCTCGAAGACCCATGCTCTTGCAATATCTGTT 2622

Db 263 GCTATGAAGCCATGCTTGGTACACTTTCCTCCAACTGTTCGGAGTACCTGGTAGAA 322

Qy 840 cgtgcttgagccaatgtctttgcctagcatagacgaattcgcctgatctcgttatcttgc 899
||| ||||| | | | | | | | | | | | | | | | |
Db 323 ACTGGCGGACAAAGTCTCATTCCATGATTAATAATGTTCGTTGGCTGCATACCTTG 382

900 tacatcgcatcattctgtttcgttcgagtttggtatttatgagtcacagagagcct 959

Db 383 T A I N T A G C A T T T T A C T A A G T A T T G T A G A G T T T G C T A T T T A T T T G G A T G C A C A G A C A C T G 442

Db 443 CACGACATAAACCGCTGTACAATATCTTCATGTAACCCATCACATTCACNATAACAG 502

Qy 1020 aatacacatctctccattgcgcgaatgattt 1051
||| ||||| || ||||| ||||| |||||
1000 aatacacatctctccattgcgcgaatgattt 1051
||| ||||| || ||||| ||||| |||||

DB 503 AACACATCTCTCCCTTNGCTGGTTGCATTT 534

RESULT	6				
AW219366					
LOCUS	AW219366	592 bp	mRNA	EST	18-MAY-2001

DEFINITION	EST301848 tomato root during/after fruit set, Cornell University
ACCESSION	AW219366
	Lycopersicon esculentum cDNA clone CLEX4M9, mRNA sequence.

```

VERSION      AW213366.1  GI:6530240
KEYWORDS
SOURCE       tomato.
EST.

```

ORGANISM Lycopodium esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE	Asteridae: euasterids I; Solanales; Solanaceae; Solanum;
AUTHORS	Lycopersicon.
TITLE	1 (bases 1 to 592)
JOURNAL	van der Hoeven,R.S., Garvin,D., Matern,A.L., Holt,I.E., Jiang,F., Upson,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Rinning,C.M., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
COMMENT	Generation of ESTs from tomato root tissue unpublished (1999) Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: http://www.genome.clemson.edu/orders/index.html 5 prime sequence.
FEATURES	Location/Qualifiers
SOURCE	1..592 /organism="Lycopersicon esculentum" /cultivar="TA496" /db_xref="taxon:4081" /clone="cLEX4M9" /clone_11b="tomato root during/after fruit set, Cornell University" /tissue_type="root" /dev_stage="plants during and after fruit-set" /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Tanksley Tissue supplied by Dave Garvin (USDA-ARS, Ithaca, NY 14850)."
BASE COUNT	152 a 125 c 125 g 190 t
ORIGIN	

Query Match	9.6%	Score 181.4;	DB 10;	Length 592;
Best Local Similarity	71.3%;	Pred. No. 7e-25;		
Matches 239;	Conservative	0;	Mismatches 96;	Indels 0;
				Gaps 0

Oy	720	gtcgttttgctcaagtcgaatctccacaaataaagcgatagtggttggaaatgtttcgg	775
Db	214	gtttatttacctaaagatggccattcccttcaaaatgacatagtgtcttggcaaatggagtt	273
Oy	780	gcaatgaagcctatgcgatcgtacacactctctctccacactgtctccgagataltgtaa	839
Db	274	gcattggaagcgtatggcattgactgctgacctccatccatcattcttgaaatgacattgaa	333
Oy	840	cgttggttgaccaaatgttttgcagatagaagaaatcgcgtgagttctgtaatttgt	899
Db	334	aattggctggactaaattgtttctgaagaaatgaagcagatgtcgatggatgacaccttgatc	393
Oy	900	tacatcgccalcatacttctgttttcgttgaattctgtatatttgaatgcagagaagcct	959
Db	394	tattgttcagttcattcgtgtaaatgtaagatttggatctatttggatgcgatcgacagctg	453
Oy	960	catgaacataagacctctctaagtaactcatcgcgcacccaataatctacaagaagcg	1011
Db	454	catgacataaactctgtgtccaaatattcttcatactgctacacatcatattatracaaacgaa	513
Oy	1020	aatacactctctccatttcgcggttaagtglttcca	1054
Db	514	aattcactctccacttctgctgttggtggcattcca	548

RESULT	7			
LOCUS	BE459711			
DEFINITION	BE459711	447 bp	mRNA	EST
ACCESSION	ESU415003	tomato developing/immature green fruit	Lycopersicon	18-MAY-2001
VERSION	BE459711	esculentum cDNA clone CLEM7H14,	mRNA sequence.	
KEYWORDS	BE459711.1	GI:9504013		
SOURCE	EST.			
ORGANISM	tomato.			
	Lycopersicon esculentum			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	Asteridae; euasterids I; Solanales; Solanaceae; Solanum;			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	SOURCE
1 (bases 1 to 447)	Alcala,J., Vredalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.B., Roming,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.	Generation of ESTs from tomato fruit tissue, immature green unpublished (2000)	Contact: CUGI	Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: http://www.genome.clemson.edu/orders/index.html 5 prime sequence.	Location/Qualifiers 1..447	/organism="Lycopersicon esculentum" /cultivar="TA496" /db_xref="taxon:4081" /clone="CLEM7H14" /clone_lib="tomato developing/immature green fruit" /tissue_type="fruit" /dev_stage="immature green (5-35 days post-anthesis)" /lab_host="SOLR" /note="Vector: pBluescriptSKMcu4apt; site_1: EcoRI; site_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and harvested at 7 day intervals through 35 dpa. Equal masses of tissue from each stage were combined (including seeds and locules) prior to mRNA isolation."
124 a	102 c	87 g	134 t			

Query Match	9.3%	Score 175.6	DB 10	Length 447
Best Local Similarity	75.9%	Pred. No. 9.2e-24		
Matches 217	Conservative 0	Mismatches 69	Indels 0	Gaps 0

Oy	1389	tttcaagagccttgatcattccaacgaagaaaggatacttcagggccgtacgcgatgata	1448
Db	159	TTTTCTGTGGTGGCAATTCACCCATTGGAGGAAATCTCACAGGCAGTGGCACACTTAATA	218
Oy	1449	gcgcgttataagtgcacaattcatctcaacaaccatatagtcctcttgatcattgaagaagc	1508
Db	219	GCTCTTTTCTTGTTGCCCTGGCATTTTACACACACACATTGGCTGTCTTTATCATTAAGAACC	278
Oy	1509	ataatgacagcgagaaatccatctgactgcattccatctgacacatcttgccattaatgggtgca	1568
Db	279	ATATGGACTCCAAATATTTCTATGACTGTATNACAAGCCAAAGGTTTGGCTTTAATGGGTGCC	338
Oy	1569	gatacacatcgaatacaccacagacatacaagcataactatgylcat.atacatactg	1628
Db	339	GTTTACCATTCACATCCACCATCTAGCTAGTACCGGCATTAATTATGGTGTCAACACATATTAGG	398
Oy	1629	atggaattggaatgtttggcctctctctaaggatcctctctttaagaagaag	1674
Db	399	ATGGACTGGATGTTTGGAACTCTTTCGCAGATCCTGTTTGAAGATGAAG	444

RESULT	8				
LOCUS	BE357415				
DEFINITION	BE357415	667 bp	mRNA	EST	20-JUL-2000
ACCESSION	DG1_15_H10.b2_A002	Dark Grown 1 (DG1)	Sorghum bicolor	cDNA, mRNA	
VERSION	BE357415				
KEYWORDS	BE357415.1	GI:9298972			
SOURCE	EST.				
ORGANISM	sorghum.				
REFERENCE	Sorghum bicolor				
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC clade; Panicoideae; Andropogoneae; Sorghum. 1 (bases 1 to 667) Cordonnier-Pirat,M.-M., Ginglé,A., Marsala,C., Sudman,M. and Pratt				

TITLE
JOURNAL
COMMENT

,L.H.
An EST database from Sorghum: dark-grown seedlings
Unpublished (2000)
Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions
below phred quality 16. The threshold for highest quality sequence
is 20.
Seq primer: JEN REV
High quality sequence stop: 593
POLYA-NO.

FEATURES
source

location/Qualifiers
1..667
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Dark Grown 1 (DG1)"
/note="Organ: 5-day-old dark-grown seedlings; Vector:
lambda zap; Site_1: XhoI; Site_2: EcoRI; The library was
made from poly-A RNA in the cloning vector lambda ZAP II.
Clones to be sequenced were prepared by mass excision."
BASE COUNT 147 a 179 c 165 g 176 t
ORIGIN

Query Match 9.28; Score 173.8; DB 10; Length 667;
Best Local Similarity 71.3%; Pred. No. 1.9e-23;
Matches 229; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

Y 734 aggtgcaattcctcaataaaggatgctgctgcaatgttggtgcaatgaagctat 793
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 293 AGATGCCATTCCCAATAGAGTAAGAGAACAAATAGTTTACATCGAAGGCTAT 352
Y 794 gccatggaactctctctcgaactgctcgcgagagatgatgaagctgtgtgaccaa 853
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 353 GCCCTTTTAACTGCTCTCCCACTTATCATATATATATGATGAGAGCGGATGACCA 412
Y 854 atgtttgcgaactgaagaatcgctgctgctgctgctgctgctgctgctgctgctg 913
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 413 GTGTACTTAAATATACAGTGAAGTGTGTTTCTATATACCTCTGTATATGCTATCTA 472
Y 914 tctgtttcgttgagcttggtgctgctgctgctgctgctgctgctgctgctgctgctg 973
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 473 TCCCATTTTGTGAGAGTTTGAATTTACTGATGCACAGACAGTGCATGACATAAAGCC 532
Y 974 tctctataaglatcgcacaccacatcatatcatcaagaagcagatacctctcc 1033
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 533 ATTATACAAATATCTGCATCAACCCATCATATTTACACAGAAATATCTTGTCTCC 592
Y 1034 attgcggtgaagtgcttca 1054
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 593 ATTTGCTGACTTGCAATTTC 613

RESULT 9
BG123510 621 bp mRNA EST 31-JAN-2001
LOCUS BG123510
DEFINITION EST469146 tomato shoot/meristem Lycopersicon esculentum cDNA clone
cTOF2K3 5' sequence similar to nearly identical to Nicotiana
tabacum sterol-C5(6)-desaturase, mRNA sequence.
ACCESSION BG123510
VERSION BG123510.1 GI:12623688
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE 1 (bases 1 to 621)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Uterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.

FEATURES
source

location/Qualifiers
1..621
/organism="Lycopersicon esculentum"
/cultivar="TA96"
/db_xref="taxon:4081"
/clone="cTOF2K3"
/clone_lib="tomato shoot/meristem"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA96). Tissue
was immediately frozen in liquid nitrogen."
BASE COUNT 163 a 137 c 126 g 195 t
ORIGIN

Query Match 9.28; Score 173.4; DB 11; Length 621;
Best Local Similarity 69.9%; Pred. No. 2.3e-23;
Matches 234; Conservative 0; Mismatches 101; Indels 0; Gaps 0;

Y 720 gtctgttcttcagatgcaatcctcaataaaggctatgctgctgcaatgttgctg 779
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 200 GTCTATATTCCTCCAAAGATGCCATACATCAAGAACATGCTGTGCAATATCTGTT 259
Y 780 gcaatgaagctatgcaatgcaatcctctctcgaactgctcgcgagatgatgaa 839
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 260 GCTATGAAGCTATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 319
Y 840 cgtgtgtggaacaaatgttgctgctgctgctgctgctgctgctgctgctgctgctg 899
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 320 AACGATGAGACCAATCTTGTGCGAATATGATGATGATGATGATGATGATGATGATG 379
Y 900 tacatgcacatctctgttctgctgctgctgctgctgctgctgctgctgctgctgctg 959
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 380 AATCGGCTATTTATCTTTGTAATGATGATGATGATGATGATGATGATGATGATGATG 439
Y 960 catgacatgaagcctctctatgaatcctcgcacacacatcatatcatcaagaagcag 1019
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 440 CATGACATTAACCTCTGTATTAATATCTGCATCTCACATCATATTTACACACAA 499
Y 1020 aatacactctcctcattgctgctgctgctgctgctgctgctgctgctgctgctg 1054
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 500 AACACACTTCCCATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 534

RESULT 10
BG598293 722 bp mRNA EST 12-APR-2001
LOCUS BG598293
DEFINITION EST496971 cSTS Solanum tuberosum cDNA clone cSTS20L7 5' sequence,
mRNA sequence.
ACCESSION BG598293
VERSION BG598293.1 GI:13616433
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 722)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chieningo,A.,
Bougl,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.

REFERENCE 1 (bases 1 to 708)
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae
; Triliceae; Hordeum.
Wing, R., Close, T. J., Kleinof, A., Wise, R., Begum, D., Frisch, D., Yu
, Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo
, T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D. W., Main, D. and
Wood, T.

TITLE Development of a genetically and physically anchored EST resource
for barley genomics
JOURNAL Unpublished (2000)
COMMENT On Aug 21, 2000 this sequence version replaced gi:9859263.
AUTHORS Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: AATTACCTCCTAAGGCG
High quality sequence start: 5
High quality sequence stop: 688.

FEATURES
source
1..708
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEH0099E02f"
/clone_1lb="Hordeum vulgare 5-45 DAP spike EST library
HVCNDA0009 (5 to 45 DAP)"
/tissue_type="5-45 DAP Spike"
/lab_host="SOLR"
/note="Vector: LambdaZAP. Site 1: EcoRI. Site 2: XhoI. For
more details on library preparation and sequence analysis
see http://www.genome.clemson.edu/projects/barley/ To
order a clone see http://www.genome.clemson.edu/orders"

BASE COUNT 144 a 211 c 182 g 167 t 4 others
ORIGIN

Query Match 8.9%; Score 167.2; DB 10; Length 708;
Best Local Similarity 71.2%; Pred. No. 3.5e-22;
Matches 220; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 734 agatgaattcctacataaaggctatgcttgcgaatgttggaatgaagctat 793
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 301 AGATGCTGTCCTACAGTGAAGCTATGAAGAAGCAATATGTCATCAAGGCTAT 360
QY 794 gcaatgtaacactctctccaaactgtctccgagatgatgaacgtgtggaacaa 853
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 361 GCCTTCTACTGCGCTCCGCTGCTGATCTGAGCATGATGAGATGATGAGACCG 420
QY 854 atgttttgctagcatagaagaattcggtggtatctgtatctgtttacatcgcatcta 913
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 421 GGTGTTTTTCATATCAAGTGAAGTGGTGGCGGATGATTTGATATGCTGCTTATA 480
QY 914 tctgttttgctgagtttggtatcttatgagatgcacagagcttcatacattaaagc 973
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 481 TCTCACCCTTGTGGAGTTTGGATTACTGATGACACAGAGATTGCATACATANAAC 540
QY 974 tctclataagatctcatccacccatcatatcatcaaaagaagaatcaactctcc 1033
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 541 ACTGTATTAAGCACCTACAGCAACCCACACATTTACAAAGAGATATCCCTATACAC 600
QY 1034 attgcgcgg 1042
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 601 ATTTGCTGG 609

RESULT 15
LOCUS BF624091 782 bp mRNA EST 21-FEB-2001
DEFINITION HVSMEH009910f Hordeum vulgare seedling shoot EST library
HVCNDA0001 (Cold stress) Hordeum vulgare cDNA clone HVSMEH009910f,

ACCESSION mRNA sequence.
VERSION BF624091.2 GI:13083024
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare

REFERENCE 1 (bases 1 to 782)
Wing, R., Close, T. J., Kleinof, A., Wise, R., Begum, D., Frisch, D., Yu
, Y., Anderson, H., Dale, J., Henry, D., Kernodle, S., Palmer, M., Rambo
, T., Saski, C., Schwartzbeck, J., Simmons, J., Choi, D. W., Main, D. and
Wood, T.

TITLE Development of a genetically and physically anchored EST resource
for barley genomics
JOURNAL Unpublished (2000)
COMMENT On Dec 18, 2000 this sequence version replaced gi:11887825.
AUTHORS Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Seq primer: AATTACCTCCTAAGGCG
High quality sequence stop: 657.

FEATURES
source
1..782
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEH009910f"
/clone_1lb="Hordeum vulgare seedling shoot EST library
HVCNDA0001 (Cold stress)"
/tissue_type="Seedling shoot"
/lab_host="TJUC121"
/note="Vector: LambdaZAP. Site 1: EcoRI. Site 2: XhoI. For
more details on library preparation and sequence analysis
see http://www.genome.clemson.edu/projects/barley/ To
order a clone see http://www.genome.clemson.edu/orders"

BASE COUNT 158 a 229 c 208 g 187 t
ORIGIN

Query Match 8.8%; Score 166.6; DB 11; Length 782;
Best Local Similarity 71.2%; Pred. No. 4.5e-22;
Matches 220; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 734 agatgaattcctacataaaggctatgcttgcgaatgttggaatgaagctat 793
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 274 AGATGCTGTCCTACAGTGAAGCTATGAAGAAGCAATATGTCATCAAGGCTAT 333
QY 794 gcaatgtaacactctctccaaactgtctccgagatgatgaacgtgtggaacaa 853
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 334 GCCTTCTACTGCGCTCCGCTGCTGATCTGAGCATGATGAGATGATGAGACCG 933
QY 854 atgttttgctagcatagaagaattcggtggtatctgtatctgtttacatcgcatcta 913
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 394 GCGTTTTTCATATCAAGTGAAGTGGTGGCGGATGATTTGATATGCTGCTTATA 453
QY 914 tctgttttgctgagtttggtatcttatgagatgcacagagcttcatacattaaagc 973
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 454 TCTCACCCTTGTGGAGTTTGGATTACTGATGACACAGAGATTGCATACATANAAC 513
QY 974 tctclataagatctcatccacccatcatatcatcaaaagaagaatcaactctcc 1033
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 514 ACTGTATTAAGCACCTACAGCAACCCACACATTTACAAAGAGATATCCCTATACAC 573
QY 1034 attgcgcgg 1042
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 574 ATTTGCTGG 582

Search completed: April 5, 2002, 04:25:21
Job time: 28616 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 03:25:38 ; Search time 89.32 Seconds
(without alignments)
4789.710 Million cell updates/sec

Title: US-09-775-879-20

Perfect score: 1889
Sequence: 1 gaagatcgatcaatcaatca.....aaatgttgatgattgtgc 1889

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 10%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	60	3.2	7218	1 US-08-232-463-14	Sequence 14, Appl
2	48	2.5	240	1 US-08-628-417-6	Sequence 6, Appl
3	46.4	2.5	1117	4 US-09-247-373B-33	Sequence 33, Appl
4	46.2	2.4	1582	3 US-08-545-196B-10	Sequence 10, Appl
5	46.2	2.4	1582	3 US-08-545-196B-12	Sequence 12, Appl
6	45.6	2.4	2223	1 US-08-257-073-4	Sequence 4, Appl
7	44.6	2.4	5394	3 US-08-688-376-1	Sequence 1, Appl
8	44.2	2.3	144	1 US-08-702-344-26	Sequence 26, Appl
9	42.8	2.3	12980	3 US-08-811-566-5	Sequence 5, Appl
10	42.2	2.2	2447	2 US-09-014-969-14	Sequence 14, Appl
11	42	2.2	124	6 5185243-1	Patent No. 5185243
12	41.8	2.2	260	2 US-08-520-678A-29	Sequence 29, Appl
13	41.8	2.2	260	4 US-08-897-126-29	Sequence 29, Appl
14	41.6	2.2	356	2 US-08-520-678A-22	Sequence 22, Appl
15	41.6	2.2	356	4 US-08-897-126-22	Sequence 22, Appl
16	41.6	2.2	9646	1 US-08-811-566-1	Sequence 1, Appl
17	41.4	2.2	140	1 US-08-628-417-5	Sequence 5, Appl
18	41.2	2.2	3527	2 US-08-909-965C-7	Sequence 7, Appl
19	41	2.2	1733	3 US-09-073-569-1	Sequence 1, Appl
20	40.8	2.2	1411	4 US-08-964-127-5	Sequence 5, Appl
21	40.8	2.2	1813	4 US-09-071-224-3	Sequence 3, Appl
22	40.6	2.1	270	2 US-08-520-678A-30	Sequence 30, Appl
23	40.6	2.1	270	4 US-08-897-126-30	Sequence 30, Appl
24	40.4	2.1	1066	1 US-08-157-101A-4	Sequence 4, Appl
25	40.4	2.1	6671	1 US-08-280-443-1	Sequence 1, Appl
26	40.4	2.1	6671	1 US-08-457-459-1	Sequence 1, Appl
27	40.4	2.1	6671	1 US-08-555-678-1	Sequence 1, Appl

28	40.4	2.1	6671	5 PCT-US95-02275-1	Sequence 1, Appl
29	40.2	2.1	361	4 US-09-018-584A-9	Sequence 9, Appl
30	40.2	2.1	1051	4 US-09-245-041-10	Sequence 10, Appl
31	40	2.1	1098	3 US-09-248-335-35	Sequence 35, Appl
32	40	2.1	1493	1 US-08-340-820-24	Sequence 24, Appl
33	40	2.1	1493	1 US-08-593-535-24	Sequence 24, Appl
34	39.6	2.1	19124	2 US-08-487-826B-13	Sequence 13, Appl
35	39.6	2.1	1882	4 US-09-370-253-1	Sequence 1, Appl
36	39.2	2.1	253	2 US-08-520-678A-25	Sequence 25, Appl
37	39.2	2.1	253	4 US-08-897-126-25	Sequence 25, Appl
38	39.2	2.1	790	4 US-09-363-970-4	Sequence 4, Appl
39	39.2	2.1	1683	4 US-09-347-803-11	Sequence 11, Appl
40	39.2	2.1	1700	2 US-08-897-340-11	Sequence 4, Appl
41	39.2	2.1	1700	2 US-08-252-329-4	Sequence 4, Appl
42	39.2	2.1	1921	2 US-08-557-128-11	Sequence 11, Appl
43	39	2.1	9589	1 US-07-925-695-1	Sequence 1, Appl
44	39	2.1	9589	1 US-07-925-695-2	Sequence 2, Appl
45	38.6	2.0	1474	4 US-08-821-994-64	Sequence 64, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
FILING DATE:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ. ID NO. 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14


```
; Sequence 10, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MONNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,196B
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FARACI, C. J.
; REGISTRATION NUMBER: 32,350
; REFERENCE/DOCKET NUMBER: 2121-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-545-196B-10

Query Match      2.4%; Score 46.2; DB 3; Length 1582;
Best Local Similarity 56.1%; Pred. No. 0.0061;
Matches 87; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 1720 ttgggtttgtctctcgtttgtctgtgtgtgttcaagttcagcttctt 1779
    || |||| || || || || || || || || || || || || || || ||
Db 1577 tttttttttttttttttttttttttttttttttttttttttttttt 1518

QY 1780 gtctcttctctctctctctatcagtctctcctcaaccttccaattatgtta 1839
    || |||| || || || || || || || || || || || || || || ||
Db 1517 tttttttttttttttttttttttttttttttttttttttttttttt 1458

QY 1840 caaacattgctgctagtttaaacatgtaatg 1874
    ||||| | || ||||| ||||
Db 1457 TAAACATATAGAAATAGAAAAAACAGTACATG 1423

RESULT 5
US-08-545-196B-12/C
; Sequence 12, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MONNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
```

```
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,196B
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FARACI, C. J.
; REGISTRATION NUMBER: 32,350
; REFERENCE/DOCKET NUMBER: 2121-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-545-196B-12

Query Match      2.4%; Score 46.2; DB 3; Length 1582;
Best Local Similarity 56.1%; Pred. No. 0.0061;
Matches 87; Conservative 0; Mismatches 68; Indels 0; Gaps 0;

QY 1720 ttgggtttgtctctcgtttgtctgtgtgtgttcaagttcagcttctt 1779
    || |||| || || || || || || || || || || || || || || ||
Db 1577 tttttttttttttttttttttttttttttttttttttttttttttt 1518

QY 1780 gtctcttctctctctctatcagtctctcctcaaccttccaattatgtta 1839
    || |||| || || || || || || || || || || || || || || ||
Db 1517 tttttttttttttttttttttttttttttttttttttttttttttt 1458

QY 1840 caaacattgctgctagtttaaacatgtaatg 1874
    ||||| | || ||||| ||||
Db 1457 TAAACATATAGAAATAGAAAAAACAGTACATG 1423

RESULT 6
US-08-257-073-4/C
; Sequence 4, Application US/08257073
; Patent No. 5765597
; GENERAL INFORMATION:
; APPLICANT: Paolletti, Enzo
; APPLICANT: de Taisne, Charles
; TITLE OF INVENTION: MALARIA RECOMBINANT POXVIRUS VACCINE
; NUMBER OF SEQUENCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue, 25th Floor
; CITY: New York
; STATE: New York
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,073
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/075,783
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

CM nucleic - nucleic search, using sw model

Run on: April 5, 2002, 03:46:03 ; Search time 221.25 Seconds

(without alignments)
7319.717 Million cell updates/sec

Title: US-09-775-879-20

Perfect score: 1889

Sequence: 1 gaagatgcatacatca.....aaatgttgatgattcttgc 1889

Scoring table: IDENTITY_NUC

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101:*

1: /SIDSL1/gcgdata/geneseq/geneseq/NA1980.DAT:*

2: /SIDSL1/gcgdata/geneseq/geneseq/NA1981.DAT:*

3: /SIDSL1/gcgdata/geneseq/geneseq/NA1982.DAT:*

4: /SIDSL1/gcgdata/geneseq/geneseq/NA1983.DAT:*

5: /SIDSL1/gcgdata/geneseq/geneseq/NA1984.DAT:*

6: /SIDSL1/gcgdata/geneseq/geneseq/NA1985.DAT:*

7: /SIDSL1/gcgdata/geneseq/geneseq/NA1986.DAT:*

8: /SIDSL1/gcgdata/geneseq/geneseq/NA1987.DAT:*

9: /SIDSL1/gcgdata/geneseq/geneseq/NA1988.DAT:*

10: /SIDSL1/gcgdata/geneseq/geneseq/NA1988.DAT:*

11: /SIDSL1/gcgdata/geneseq/geneseq/NA1990.DAT:*

12: /SIDSL1/gcgdata/geneseq/geneseq/NA1991.DAT:*

13: /SIDSL1/gcgdata/geneseq/geneseq/NA1992.DAT:*

14: /SIDSL1/gcgdata/geneseq/geneseq/NA1993.DAT:*

15: /SIDSL1/gcgdata/geneseq/geneseq/NA1994.DAT:*

16: /SIDSL1/gcgdata/geneseq/geneseq/NA1995.DAT:*

17: /SIDSL1/gcgdata/geneseq/geneseq/NA1996.DAT:*

18: /SIDSL1/gcgdata/geneseq/geneseq/NA1997.DAT:*

19: /SIDSL1/gcgdata/geneseq/geneseq/NA1998.DAT:*

20: /SIDSL1/gcgdata/geneseq/geneseq/NA1999.DAT:*

21: /SIDSL1/gcgdata/geneseq/geneseq/NA2000.DAT:*

22: /SIDSL1/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	489.8	25.9	1164	AAAC41897	Arabidopsis thaliana
2	164.2	8.7	1381	AAAH0964	Sterol C3 desaturase
3	158.2	8.4	446	AAAF9689	Eucalyptus grandis
4	134.6	7.1	936	AAAF8252	Oligonucleotide D1
5	134.6	7.1	936	AAAF8254	Oligonucleotide D1
6	134.6	7.1	936	AAAF8257	Oligonucleotide D1
7	134.6	7.1	936	AAAF8259	Oligonucleotide D1
8	134.6	7.1	936	AAAF8262	Oligonucleotide D2
9	134.6	7.1	936	AAAF8265	Oligonucleotide D1
10	133.8	7.1	936	AAAF8252	Oligonucleotide D1
11	133.8	7.1	936	AAAF8254	Oligonucleotide D1

12	133.8	7.1	936	AAAF8257	Oligonucleotide D1
13	133.8	7.1	936	AAAF8259	Oligonucleotide D2
14	133.8	7.1	936	AAAF8262	Oligonucleotide D2
15	133.8	7.1	938	AAAF8255	Oligonucleotide D1
16	54.2	2.9	1954	AAAZ3441	cDNA encoding huma
17	52.6	2.8	2158	AAAF16153	Human prostate can
18	49.8	2.6	7372	AAAX3181	Base sequence of t
19	49.8	2.6	7372	AAAX3182	Base sequence of t
20	49.8	2.6	7797	AAAX3180	Cowpox virus bsr f
21	49.8	2.6	7996	AAAX3184	Base sequence of t
22	49.6	2.6	4590	AAH24065	Yeast AOD9604-asso
23	49.4	2.6	244	AAAF58238	Oligonucleotide D1
24	49.2	2.6	244	AAAF58238	Oligonucleotide D1
25	48.6	2.5	2338	AAAO54656	Plasmid pNX1 comp
26	48	2.5	240	AAAT76782	Staphylococcus aur
27	48	2.5	2195	AAAI4072	Human SPORU-1 pa
28	48	2.5	2364	AAZ33323	Human secreted pro
29	47.2	2.5	1020	AAZ97182	Human prostate can
30	47.2	2.5	3375	AAZ07192	Human lung tumour
31	47.2	2.5	3275	AAAC79145	Human lung tumour
32	47	2.5	2921	AAAC7829	Human cancer assoc
33	46.8	2.5	268	AAH70080	Human cervical can
34	46.8	2.5	2057	AAAF72803	Secreted protein g
35	46.4	2.5	1117	AAAF59475	Nucleotide sequenc
36	46.4	2.5	3060	AAAD02924	Human PRO5723 cDNA
37	46.4	2.5	3060	AAAF44263	Human PRO5723 nucl
38	46.4	2.5	3060	AAAC91489	Human PRO5723 cDNA
39	46.2	2.4	1582	AAAT28259	Survival motor neu
40	46.2	2.4	1582	AAAT1831	Human survival mot
41	46	2.4	464	AAH33241	Human colon cancer
42	46	2.4	1119	AAAC60033	Human secreted pro
43	46	2.4	1492	AAAC98102	Human colon cancer
44	46	2.4	1493	AAAH34433	Human colon cancer
45	46	2.4	2475	AAAT77667	Human PRO1246 cDNA

ALIGNMENTS

RESULT 1

AAAC41897 standard; DNA: 1164 BP.

AC AAAC41897;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 33540.

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PF 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-01231825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154019.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.

Db 785 726
 Cy 448 tcttggtgacgtacgaatttgatttgatgtgacatgctgattgattatg 507
 Db 725 666
 Cy 508 attatttgcatcttcacatcaaggaattgaacagttagtgcttataaatttt 567
 Db 665 606
 Cy 568 gtgcaacaatgagaagtcgtacatcttgaaattgaatttctacttgcaattaaagtc 627
 Db 605 546
 Cy 628 caactaaattgttgtagaagattgtctacttccagacaactcttctgctctcc 687
 Db 545 486
 Cy 688 tgaagactgctctagttgaaatcttlttgctgttctgctcagatgcattccta 747
 Db 485 426
 Cy 748 caataaagctatgcgtttgcaaatgttgygcaatgaagctatgcattgacactc 807
 Db 425 366
 Cy 808 tcttcacaactgctccgagagatgatggaacgtggttgagccaaatttctgtagca 867
 Db 365 306
 Cy 868 tagagcaatcgctgagatctctgatttgtttacatcgcatctatctgttctgtg 927
 Db 305 246
 Cy 928 agtttgatattatgtgacagagagctcatgacattaaagctctataagtlac 987
 Db 245 186
 Cy 988 tccatgacacccatcatatctacacaagcagaatacactctctcattgcccgaagt 1047
 Db 185 126
 Cy 1048 gtcttcagttgtctctcttagtctctgttaaaagattgtagcattttagtcttaca 1107
 Db 125 66
 Cy 1108 gaaaagacttgtagcagcgtgctgtactccaatcacatttgcattcttaccata 1167
 Db 65 6
 Cy 1168 aagta 1172
 Db 5 1
 RESULT 5
 AAF58254/C
 ID AAF58254 standard: DNA; 936 BP.
 AC AAF58254;
 XX 24-APR-2001 (first entry)
 DE Oligonucleotide D1875.
 XX
 XX Eletion-transfer group; ETM; mismatch; genotyping;
 KM gene expression; ss.
 CS Synthetic.
 XX
 XX W0200107665-A2.
 FN
 XX

PD 01-FEB-2001.
 XX
 XX 26-JUL-2000; 2000MO-US20476.
 XX
 XX 26-JUL-1999; 99US-0145695.
 PR 17-MAR-2000; 2000US-0190259.
 XX
 XX (CLIN-) CLINICAL MICRO SENSORS INC.
 XX
 XX Umek RM;
 PI
 XX WPI; 2001-159728/16.
 DR
 XX
 XX Nucleic acids containing electron-transfer group, useful as labels in
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
 PT a single surface
 XX
 XX Example 6; Page 127; 159pp; English.
 PS
 CC The present invention relates to a composition comprising two nucleic
 CC acids each containing an electron-transfer group (ETM) having
 CC different redox potentials. The invention is used for electronic
 CC detection of nucleic acids, especially of substitutions (mismatches)
 CC and single-nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.
 XX
 SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;
 Query Match 7.1%; Score 134.6; DB 22; Length 936;
 Best Local Similarity 0.6%; Pred. No. 6, 1e-24;
 Matches 5; Conservative 498; Mismatches 282; Indels 0; Gaps 0;
 QY 388 tattaactatgcttaactgcttctctatgtatcgatttccaatttaagagaggt 447
 Db 785 726
 QY 448 tcttggtgacgtacgaatttgatttgatgtgacatgctgattgattatg 507
 Db 725 666
 QY 508 attatttgcatcttcacatcaaggaattgaacagttagtgcttataaatttt 567
 Db 665 606
 QY 568 gtgcaacaatgagaagtcgtacatcttgaaattgaatttctacttgcaattaaagtc 627
 Db 605 546
 QY 628 caactaaattgttgtagaagattgtctacttccagacaactcttctgctctcc 687
 Db 545 486
 QY 688 tgaagactgctctagttgaaatcttlttgctgttctgctcagatgcattccta 747
 Db 485 426
 QY 748 caataaagctatgcgtttgcaaatgttgygcaatgaagctatgcattgacactc 807
 Db 425 366
 QY 808 tcttcacaactgctccgagagatgatggaacgtggttgagccaaatttctgtagca 867
 Db 365 306
 QY 868 tagagcaatcgctgagatctctgatttgtttacatcgcatctatctgttctgtg 927
 Db 305 246
 QY 928 agtttgatattatgtgacagagagctcatgacattaaagctctataagtlac 987
 Db 245 186

```
OY 988 tcatgccaccacatcatctacacaagcagaatcacctctccattgcggtagt 1047
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 185 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 126
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1048 gtcttcagttgtctctctttagtcttgtaagaattgtagcatttagttcttaca 1107
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 125 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 66
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1108 gaaagaactgttcagagactgtctgtactccaaatcacattgtcattccttaccata 1167
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 65 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 6
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1168 aagta 1172
      : : : :
DB 5 wwwwww 1

RESULT 6
AAF58257/c
ID AAF58257 standard; DNA; 936 BP.
XX
AC AAF58257;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1954.
XX
KM Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PE 26-JUL-2000; 2000MO-US20476.
XX
PR 26-JUL-1999; 9905-0145695.
XX
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface.
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match 7.1%; Score 134.6; DB 22; Length 936;
Best Local Similarity 0.6%; Pred. No. 6.1e-24;
Matches 5; Conservative 499; Mismatches 282; Indels 0; Gaps 0;
```

```
DB 725 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 666
OY 508 atattgtgcatactccactcaaggagatggaacattgagcttataaagttct 567
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 665 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 606
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 568 gtgcaacaaatgagaagtcgtacacttgaattgaatttctacttgccattaaagtc 627
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 605 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 546
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 628 cacttaaatgttgtaggaagtgtgtcactcttcagacacattcttctgctctc 687
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 545 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 486
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 688 tgaactctgtcttagttgaacatttttggctgtttgtctcagaatgcata 747
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 485 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 426
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 748 caataaaggctatgcttgcgaatgttgcgaatggaagctatgcattgacac 807
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 425 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 366
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 808 ttcttcaactgtctccgagagatgtatgaagtggttgaccacaaatgtttgtcga 867
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 365 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 306
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 868 tagacgaattcgctgattctgtatttcttaccatgcacatcttctgtctg 927
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 305 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 246
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 928 agtttgatattatgtatgcacagagacttcatgaatgaagcctctcataagtc 987
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 245 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 186
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 988 tcatgccaccacatcatctacacaagaagaatacactctccattgcggtagt 1047
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 185 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 126
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1048 gtcttcagttgtctctctttagtcttgtaagaattgtagcatttagttcttaca 1107
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 125 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 66
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1108 gaaagaactgttcagagactgtctgtactccaaatcacattgtcattccttaccata 1167
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 65 wwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwwww 6
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 1168 aagta 1172
      : : : :
DB 5 wwwwww 1

RESULT 7
AAF58259/c
ID AAF58259 standard; DNA; 936 BP.
XX
AC AAF58259;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D2004.
XX
KM Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PE 26-JUL-2000; 2000MO-US20476.
XX
PR 26-JUL-1999; 9905-0145695.
```

PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XX
XX
PS Example 6; Page 128; 159pp; English.
XX
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
XX monitoring gene expression.
XX
XX Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;
XX

Query Match	7.1%;	Score 134.6;	DB 22;	Length 936;
Best Local Similarity	0.6%;	Pred. No. 6.1e-24;		
Matches	5;	Conservative 498;	Mismatches 282;	Indels 0;
				Gaps 0;

[illegible]

RESULT	8
AAF58262/c	
ID	AAF58262 standard; DNA; 936 BP.

AC AAF58262;

DT 24-APR-2001 (first entry)

DE Oligonucleotide D2007.

KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss

XX
XX

XX WO20010766
PNXX 01-FEB-2001
PD

XX
PF 26-JUL-2000: 2000WO-US20476.

XX 26-JUL-1999; 99US-0145695.
PR

PK 1 / - MARK - 2000; 2000005-0190239.
XX

CLINICAL MICRO SENSORS INC
(CEN)
XX
XX

XX

3	XX		
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			
15			
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			
61			
62			
63			
64			
65			
66			
67			
68			
69			
70			
71			
72			
73			
74			
75			
76			
77			
78			
79			
80			
81			
82			
83			
84			
85			
86			
87			
88			
89			
90			
91			
92			
93			
94			
95			
96			
97			
98			
99			
100			

PT hybridization assays, e.g. for genotyping, allowing repeat analyses on

[illegible]

xx
xx
ffh
presen+
; nren+;
[a+c
+c
c
com
;+;
co
3;
v;
r
+;
f:
]o;

CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic

cc detection of nucleic acids, especially of substitution;
cc and single-nucleotide polymorphisms, e.g. for genotyping

CC monitoring gene expression.
XX

sequence 530 BP; 3 A; 139 C; 10 G; 8 T; 116 Other;

Query Match	7.1%	Score 134.6	DB 22	Length	936
Best Local Similarity	0.6%	Pred. No. 6.1e-24			
Matches	5	Conservative	498	Mismatches	282
				Indels	0
				Gaps	0

388 tattcaactatgtctaatcgctttctatgttatcgatttttcaattcaa1gaagaggtt 447

[illegible]

QY 448 tcttcgtgtactgttacagtaatttggatlttgatgtgatagtltcatgtt:gcatttattg 507

[illegible]

508 a l c a l l c g c a r a l c c c c a l c c c a g g a t c g a a c a g t t a g t c g c t t a : a t c a a g c t t t t 567

[illegible]


```
Db 65 WWWWWW
QY 1168 aagta 1172
Db 5 WWWWW 1

RESULT 10
AAF58252
ID AAF58252 standard; DNA; 936 BP.
XX
AC AAF58252;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1875.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KM gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PE 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SO Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other:

Query Match 7.1%; Score 133.8; DB 22; Length 936;
Best Local Similarity 0.8%; Pred. No. 9,7e-24;
Matches 6; Conservative 493; Mismatches 280; Indels 0; Gaps 0;

QY 326 ttctacatcttacccttaaaatcaagcttaacctcccaaggctcgaacttcaactt 385
Db 3 WWWWWW
QY 386 tgtatcacctatgtctatgccttcctatgctatgctatcttcaatgaaggagag 445
Db 63 WWWWWW
QY 446 ttctctgctagctgaagtaattgattgattgattgattgattgattgattgatt 505
Db 123 WWWWWW
QY 506 tgaattattgctacatactcactaaggaattgaaagattgattgattgattgatt 565
Db 183 WWWWWW
242
```

```
QY 566 ttgtcaccaatgagaagctgacatcttgaattgaatttctactgcatcattaa 625
Db 243 WWWWWW
QY 626 tccacttaattgttgtgaagtgatgctacttcaagacatcttctgcttc 685
Db 303 WWWWWW
QY 686 tctgagactcgtcttaattgaaactttttgctcgttttgccttaaatgcaatcc 745
Db 363 WWWWWW
QY 746 tacaataagctatgcttgcaaatgttgtgcaatgaagctatgcatgtacac 805
Db 423 WWWWWW
QY 806 tcttcccaactgctccgagatgattgaacgttgttgaccaaagtgtgctag 865
Db 483 WWWWWW
QY 866 catagaagcaatcgctgattcgtattgtttacatgcacatctatgttttctg 925
Db 543 WWWWWW
QY 926 tgaatttgattatttgatgacagagagcttcaatgaagctctctaatgta 985
Db 603 WWWWWW
QY 986 tctcatgcccacacatcatctacaagaagatacactctcacttgcggtaa 1045
Db 663 WWWWWW
QY 1046 gtgtttcaagttgtctctctttagtctctgtaaaagattgtagcaattgttctta 1104
Db 723 WWWWWW
781

RESULT 11
AAF58254
ID AAF58254 standard; DNA; 936 BP.
XX
AC AAF58254;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1875.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KM gene expression; ss.
XX
OS Synthetic.
XX
PN WO200107665-A2.
XX
PD 01-FEB-2001.
XX
PE 26-JUL-2000; 2000WO-US20476.
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
DR WPI; 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
PS Example 6; Page 127; 159pp; English.
XX
```


[illegible]

Query Match	7.1%;	Score 133.8;	DB 22;	Length 936;
Best Local Similarity	0.88;	Pred. No. 9.7e-24;		
Matches	6;	Conservative 493;	Mismatches 280;	Indels 0;
				Gaps 0

[illegible]

RESULT	14
AAFS58262	
ID	AAFS58262 standard; DNA, 936 BP.
XX	
AC	AAFS58262;
XX	
DT	24-APR-2001 (first entry)
XX	
DE	Oligonucleotide D2007.
XX	
KW	Electron-transfer group; ETM: mismatch; genotyping;
KW	gene expression; ss.
XX	
OS	Synthetic.
XX	
PN	WO200107665-A2.
XX	
PD	01-FEB-2001.
XX	

PF 26-JUL-2000; 2000MO-US20476.
XX
XX 26-JUL-1999; 9905-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
DR WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface
XX
XX Example 6; Page 128; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX
SQ Sequence 936 BP; 5 A; 139 C; 10 G; 6 T; 776 other;

Query Match 7.1%; Score 133.8; DB 22; Length 936;
Best Local Similarity 0.8%; Pred. No. 9.7e-24;
Matches 6; Conservative 493; Mismatches 280; Indels 0; Gaps 0;

OY 326 ttctacatctatctaaatacaacgtttacctcccaaggctcgcacttactt 385
DB 3 www.
OY 386 tgaattcaactgtcttaacgtcttctctatgtatgtatttcaatttaagaagg 445
DB 63 www.
OY 446 ttcttcgttactgtacagtaattgtatgtatgtatgtatgtatgtatgtat 505
DB 123 www.
OY 506 tgattattgtgcatctccatcctaaggatgaacagttagctgctataaagtt 565
DB 183 www.
OY 566 ttgtgcaaccaaagagtcgtacatcttgaagtgtgaatttctactgtccattaa 625
DB 243 www.
OY 626 tccacttaattgttgttgaagtgatgtctacttcaagacatcttcttctgtc 685
DB 303 www.
OY 686 tctgagactgtctgttgaatcttcttgtctgttctgtctcagaatgcattcc 745
DB 363 www.
OY 746 taacataaaggctatgtgttgcataatgttggcaatgaagctatgcattgtacac 805
DB 423 www.
OY 806 tcttctccaactgtcttcgagatgatgtgaacgtgttgagcaaatgtttgtctag 865
DB 483 www.
OY 866 catagacgaattcggctggtatctgtatttgttacaatgcatacttgttctgt 925
DB 543 www.
OY 926 tgaattgttatctatgtgctgacagagctcattgaacttaagctctctataagta 985

DB 603 www.
OY 986 ttccatgcccacccatcatatctacaagaacgaatacacatcttccatttgcggtaa 1045
DB 663 www.
OY 1046 gtgtttcgaattgtcttctttagtcttctgtlaaagaatgtagcatatttctta 1104
DB 723 www.

RESULT 15

AAF58255 standard; DNA; 938 BP.

AAF58255;

24-APR-2001 (first entry)

Oligonucleotide D1876.

Electron-transfer group; ETM; mismatch; genotyping;

gene expression; ss.

Synthetic.

WO200107665-A2.

01-FEB-2001.

26-JUL-2000; 2000MO-US20476.

26-JUL-1999; 9905-0145695.

17-MAR-2000; 2000US-0190259.

(CLIN-) CLINICAL MICRO SENSORS INC.

Umek RM;

WPI; 2001-159728/16.

Nucleic acids containing electron-transfer group, useful as labels in

hybridization assays, e.g. for genotyping, allowing repeat analyses on

a single surface

Example 6; Page 127; 159pp; English.

The present invention relates to a composition comprising two nucleic

acids each containing an electron-transfer group (ETM) having

different redox potentials. The invention is used for electronic

detection of nucleic acids, especially of substitutions (mismatches)

and single-nucleotide polymorphisms, e.g. for genotyping,

monitoring gene expression.

Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match 7.1%; Score 133.8; DB 22; Length 938;

Best Local Similarity 0.8%; Pred. No. 9.7e-24;

Matches 6; Conservative 493; Mismatches 280; Indels 0; Gaps 0;

OY 326 ttctacatctatctaaatacaacgtttacctcccaaggctcgcacttactt 385
DB 3 www.
OY 386 tgaattcaactgtcttaacgtcttctctatgtatgtatgtatgtatgtatgtat 445
DB 63 www.
OY 446 ttcttcgttactgtacagtaattgtatgtatgtatgtatgtatgtatgtat 505
DB 123 www.

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 17:06:42 ; Search time 23.59 Seconds
(without alignments)
1742.371 Million cell updates/sec

Title: US-09-775-879-21

Perfect score: 1562
Sequence: 1 MAADNAVYLMQFVDETSFYNR.....SLRDPLEEDONKDSFKRAE 281

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL_17:*
2: sp-archaea:*
3: sp-bacteria:*
4: sp-fungi:*
5: sp-human:*
6: sp_invertebrate:*
7: sp_mhcl:*
8: sp-organelle:*
9: sp-phage:*
10: sp-plant:*
11: sp-rodent:*
12: sp-virus:*
13: sp-vertebrate:*
14: sp-unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1562	100.0	281	10	09SWU4 arabidopsi
2	1555	99.6	281	10	09SYS2 arabidopsi
3	1538	98.5	281	10	039208 arabidopsi
4	1265.5	81.0	279	10	09M883 arabidopsi
5	1199.5	76.8	271	10	09ZT29 arabidopsi
6	1165	74.6	271	10	09XE12 arabidopsi
7	359	23.0	299	11	09E0S5 ratu
8	330	21.1	299	4	09UK15 ratu
9	315.5	20.2	302	5	09U174 ratu
10	297	19.0	386	3	093875 leishman
11	279.5	17.9	300	3	094457 schizosac
12	176	11.3	261	10	09LNM2 arabidopsi
13	174.5	11.2	286	5	020612 caenorhabd
14	171	10.9	115	3	09U010 schizosac
15	167	10.7	253	10	09ZM22 arabidopsi
16	164.5	10.5	274	2	09A7P9 caulobacter
17	160.5	10.3	274	5	09BHE9 leishman
18	152.5	9.8	231	2	055695 synechocyst
19	152	9.7	293	11	09CRA4 mus musculi

20	149.5	9.6	258	10	09FW22 arabidopsi
21	149	9.5	300	5	020027 caenorhabd
22	148.5	9.5	293	11	035532 ratu
23	147.5	9.4	269	5	020611 caenorhabd
24	147.5	9.4	278	5	09VYA2 drosophila
25	146	9.3	300	3	09U0H4 schizosac
26	141.5	9.1	567	10	039045 arabidopsi
27	138.5	8.9	580	10	023679 arabidopsi
28	138	8.8	303	10	09S276 arabidopsi
29	138	8.8	406	5	09VID2 drosophila
30	137	8.8	302	4	09BSX6 homo sapien
31	135.5	8.7	625	10	022681 arabidopsi
32	135	8.6	333	6	09GKT2 macaca fasc
33	135	8.6	333	6	09BG04 macaca fasc
34	129.5	8.3	258	10	09M9T3 arabidopsi
35	129.5	8.3	259	10	09AST3 arabidopsi
36	127	8.1	820	10	049656 arabidopsi
37	122.5	7.8	604	10	023678 arabidopsi
38	122	7.8	255	2	0912K0 pseudomonas
39	120.5	7.7	622	10	039046 arabidopsi
40	119.5	7.7	621	10	09PRK6 oryza sativ
41	118	7.6	623	10	039047 arabidopsi
42	117	7.5	293	3	094298 schizosac
43	115.5	7.4	621	10	09XH51 oryza sativ
44	115.5	7.4	635	10	080938 arabidopsi
45	115	7.4	246	2	P72982 synechocyst

ALIGNMENTS

RESULT 1
ID 09SWU4 PRELIMINARY: PRT: 281 AA.
AC 09SWU4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE DELTA7 STEROL C-5 DESATURASE.
GN STEL.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids; II: Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. WS-2 (MASSILEMSKIDA-2);
RX MEDLINE=99128169; PubMed=9927639;
RA Choe S., Noguchi T., Fujioaka S., Takatsuto S., Tissier C.P.,
RA Feldmann K.A.;
RT "The Arabidopsis dwf7/stel mutant is defective in the delta7 sterol C-5 desaturation step leading to brassinosteroid biosynthesis.";
RL Plant Cell 11:207-221(1999).
DR EMBL: AF105034; AAD38120.1; -
DR InterPro: IPR001541; Sterol-desat.
DR Pfam: PF01598; Sterol-desat; 1.
FT VARIANT 60 W -> *.
FT VARIANT 230 W -> *.
SQ SEQUENCE 281 AA: 33201 MW: ED1A53AEF394652A CRC64.

Query Match 100.0%; Score 1562; DB 10; Length 281;
Best local similarity 100.0%; Pred. No. 2.8e-128;
Matches 281; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAADNAVYLMQFVDETSFYNRIVLSLLPANLMEPLPHRLQTWLKNYLA;TLTYISGFLW 60
DB 1 MAADNAVYLMQFVDETSFYNRIVLSLLPANLMEPLPHRLQTWLKNYLA;TLTYISGFLW 60
QY CFYIYVLTAKINYLKPKDAIPTIKARLQMFVAMKAMPWTLLPTVSESN;ERGMKCFASI 120
|||||

```
Db 61 CFYIYKLNIVLPKDAITIKAMRLQMFVANKAMPWYLLPTVSESMERGTCFASI 120
OY 121 DEFGWILYFVYIAYLVFEVFGIYWMHRELHDIKPLKYKLTATHTHYNKNTLSPFAGLA 180
Db 121 DEFGWILYFVYIAYLVFEVFGIYWMHRELHDIKPLKYKLTATHTHYNKNTLSPFAGLA 180
OY 181 FHPVDGIILOAVPHVIALFTVPIHFTTHIGLLFMEAIWTANIHDCIHGNTIPWGAGYHTI 240
Db 181 FHPVDGIILOAVPHVIALFTVPIHFTTHIGLLFMEAIWTANIHDCIHGNTIPWGAGYHTI 240
OY 241 HHTTYKHNGHYTIWMDMFGSLRDLLEEDNKSFKAE 281
Db 241 HHTTYKHNGHYTIWMDMFGSLRDLLEEDNKSFKAE 281

RESULT 2
O9SYS2 PRELIMINARY: PRT: 281 AA.
AC O9SYS2:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE STEROL-C5-DESATURASE.
GN F16B3.21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA; TISSUE=LEAVES;
RX MEDLINE=99273992; PubMed=10344195;
RA Huseinstein T., Schaller H., Gachotte D., Benveniste P.;
RT "Delta-sterol-C5-desaturase:molecular characterization and functional
RT expression of wild-type and mutant alleles.";
RL Plant Mol. Biol. 39:891-906(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
RA Ronging C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Nierman W.C., Fraser J.C.M.;
RT "Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence.";
RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF069468; AAD12944.1; -.
DR EMBL; AC021640; AAF32465.1; -.
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
SQ SEQUENCE 281 AA: 33142 MW: D32B1BCDC467D7 CRC64:

Query Match 99.6%: Score 1555; DB 10: Length 281;
Best Local Similarity 99.6%: Pred No. 1.1e-127;
Matches 280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Db 241 HHTTYKHNGHYTIWMDMFGSLRDLLEEDNKSFKAE 281
|||||
RESULT 3
O9SYS2 PRELIMINARY: PRT: 281 AA.
AC O9SYS2:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE STEROL-C5-DESATURASE.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=97077348; PubMed=8919915;
RX MEDLINE=97077348; PubMed=8919915;
RA Gachotte D., Huseinstein T., Bard M., Lacroste F., Benveniste P.;
RT "Isolation and characterization of an Arabidopsis thaliana cDNA
RT encoding a delta 7-sterol-C5-desaturase by functional complementation
RT of a defective yeast mutant.";
RL Plant J. 9:391-398(1996).
DR EMBL; X90454; CAA62079.1; -.
DR Mendel; 12510; Arabid.1710.12510.
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
SQ SEQUENCE 281 AA: 33121 MW: 84C7E6167D2974B7 CRC64:
```

```
Query Match 98.5%: Score 1538; DB 10: Length 281;
Best Local Similarity 98.6%: Pred No. 3.4e-126;
Matches 277; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 MAADNAVLMQFVDETSFYNRIVLSHLLPANLWEPRLPHLOTWLRNLAAGTLTYISGFILM 60
Db 1 MAADNAVLMQFVDETSFYNRIVLSHLLPANLWEPRLPHLOTWLRNLAAGTLTYISGFILM 60
OY 61 CFYIYKLNIVLPKDAITIKAMRLQMFVANKAMPWYLLPTVSESMERGTCFASI 120
Db 61 CFYIYKLNIVLPKDAITIKAMRLQMFVANKAMPWYLLPTVSESMERGTCFASI 120
OY 121 DEFGWILYFVYIAYLVFEVFGIYWMHRELHDIKPLKYKLTATHTHYNKNTLSPFAGLA 180
Db 121 DEFGWILYFVYIAYLVFEVFGIYWMHRELHDIKPLKYKLTATHTHYNKNTLSPFAGLA 180
OY 181 FHPVDGIILOAVPHVIALFTVPIHFTTHIGLLFMEAIWTANIHDCIHGNTIPWGAGYHTI 240
Db 181 FHPVDGIILOAVPHVIALFTVPIHFTTHIGLLFMEAIWTANIHDCIHGNTIPWGAGYHTI 240
OY 241 HHTTYKHNGHYTIWMDMFGSLRDLLEEDNKSFKAE 281
Db 241 HHTTYKHNGHYTIWMDMFGSLRDLLEEDNKSFKAE 281

RESULT 4
O9M883 PRELIMINARY: PRT: 279 AA.
AC O9M883:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE PUTATIVE STEROL-C5-DESATURASE.
GN F16B3.22.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
```

RC STRAIN-CV, COLUMBIA;
3A Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B., Wu D.,
3B Rongling C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
3A Bowman C.L., White O., Niernan W.C., Fraser C.M.,
3T "Arabidopsis thaliana chromosome III BAC F1683 genomic sequence";
3L Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC021640; AAF32466.1;
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
SQ SEQUENCE 279 AA; 3311 MW; DCAB107445FA2795 CRC64;

Query Match 81.0%; Score 1265.5; DB 10; Length 279;
Best Local Similarity 79.4%; Pred. No. 1,7e-102;
Matches 224; Conservative 26; Mismatches 27; Indels 5; Gaps 2;

QY 1 MAADNA-YLMQFVDETSFYNRIVLSHLLPANLMEPLPHFLOTWLRNYLAGTLLYFISGFL 59
DB 1 MAATMADYNDQIVNETSFYNNRMVLSHLLPVNLMEPLPHFLOTWLRNYLAGTLLYFISGFL 60
QY 60 WCFYIYKINIVLPKAIPTIKAMRLQMFVAMKAMPYTLPTVSESMIERGWTCKFAS 119
DB 61 WCFYIYKINIVLPKAIPTIKAMRLQMFVAMKAMPYTLPTVSESMIERGWTCKFAS 120
QY 120 IDEFGWILYFYIAIVYEFEGIYMMHRELHDIKPLKYLAHTHHIYNKONTLSPEFAGL 179
DB 121 LDHNMFLCFIYALYLVLEFMYVWKELHDIKPLKHLATHHHIYNKONTLSPEFAGL 180
QY 180 AFHPVDGILQAVPHVIALFPIVPIHTTHIGLLEMEAIWTANIHDCIHGNIWPMVAGAYT 239
DB 181 AFHPLDGLQAVPHVIALFPIVPIHTTHIGLLEMEAIWTANIHDCIHGNIWPMVAGAYT 240
QY 240 IHHTYKHNHGYTIWMDMFGSLRDLPLEDDDKDSFKKAE 281
DB 241 IHHTYKHNHGYTIWMDMFGSLRDLPLEDDDKDSFKKAE 282

RESULT 5
092T29 PRELIMINARY; PRT; 271 AA.
AC 092T29;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SPROL-C5(6)-DESATURASE.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, XANTHI SH6; TISSUE-CALLI DERIVED FROM LEAF PROTOPLASTS;
RA Huseinstein T., Schaller H., Gachotte D., Benveniste P.;
RT "sterol-C5(6)-desaturase: molecular characterization and functional
expression of wild-type and mutant alleles";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF081194; AAD04034.1;
DR Mendel: 36287; Nica:1710;36287.
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
SQ SEQUENCE 271 AA; 31842 MW; 361520EAAB56D86F CRC64;

Query Match 76.8%; Score 1199.5; DB 10; Length 271;
Best Local Similarity 77.2%; Pred. No. 9.2e-97;
Matches 207; Conservative 27; Mismatches 33; Indels 1; Gaps 1;

QY 7 YLMQFVDETSFYNRIVLSHLLPANLMEPLPHFLOTWLRNYLAGTLLYFISGFLCFYIY 66
DB 4 YLMQFVDETSFYNRIVLSHLLPANLMEPLPHFLOTWLRNYLAGTLLYFISGFLCFYIY 63
QY 67 LKINIVLPKDAIPTIKAMRLQMFVAMKAMPYTLPTVSESMIERGWTCKFASIDEFGMI 126

DB 64 LKRNVIKDAIPSKAKMLQISVAMKAMPYCALPSLSEYVIEGWTCKFARISDVGM 123
QY 127 LYEVYIYLVFEFGIYMMHRELHDIKPLKYLAHTHHIYNKONTLSPEFAGLAFHPVDG 186
DB 124 SYVIAAYLVYVERGIIYMMHRELHDIKPLKYLAHTHHIYNKONTLSPEFAGLAFHPVDG 183
QY 187 ILQAVPHVIALFPIVPIHTTHIGLLEMEAIWTANIHDCIHGNIWPMVAGAYTTHHTTYK 246
DB 184 ILQAVPHVIALFPIVPIHTTHIGLLEMEAIWTANIHDCIHGNIWPMVAGAYTTHHTTYK 243
QY 247 HNYGHYTIWMDMFGSLRDLPLEDDDK 274
DB 244 HNYGHYTIWMDMFGSLRDLPLEDDDK 270

RESULT 6
09XE12 PRELIMINARY; PRT; 271 AA.
AC 09XE12;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE SPROL-C5(6)-DESATURASE HOMOLOG.
OS Nicotiana tabacum (Common tobacco).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX NCBI_TaxID=4097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, XANTHI SH6; TISSUE-LEAF;
RA Huseinstein T., Schaller H., Gachotte D., Benveniste P.;
RT "Delta7-sterol-C5-desaturase: molecular characterization and
functional expression of wild-type and mutant alleles";
RL Plant Mol. Biol. 39:891-906(1999).
DR EMBL: AF099669; AAD04058.1;
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
SQ SEQUENCE 271 AA; 32036 MW; 1E705A0B69C320D4 CRC64;

Query Match 74.6%; Score 1165; DB 10; Length 271;
Best Local Similarity 75.8%; Pred. No. 9.3e-94;
Matches 200; Conservative 29; Mismatches 35; Indels 0; Gaps 0;

QY 7 YLMQFVDETSFYNRIVLSHLLPANLMEPLPHFLOTWLRNYLAGTLLYFISGFLCFYIY 66
DB 4 YLMQFVDETSFYNRIVLSHLLPANLMEPLPHFLOTWLRNYLAGTLLYFISGFLCFYIY 63
QY 67 LKINIVLPKDAIPTIKAMRLQMFVAMKAMPYTLPTVSESMIERGWTCKFASIDEFGMI 126
DB 64 LKRNVIKDAIPSKAKMLQISVAMKAMPYCALPSLSEYVIEGWTCKFARISDVGM 123
QY 127 LYEVYIYLVFEFGIYMMHRELHDIKPLKYLAHTHHIYNKONTLSPEFAGLAFHPVDG 186
DB 124 SYVIAAYLVYVERGIIYMMHRELHDIKPLKYLAHTHHIYNKONTLSPEFAGLAFHPVDG 183
QY 187 ILQAVPHVIALFPIVPIHTTHIGLLEMEAIWTANIHDCIHGNIWPMVAGAYTTHHTTYK 246
DB 184 ILQAVPHVIALFPIVPIHTTHIGLLEMEAIWTANIHDCIHGNIWPMVAGAYTTHHTTYK 243
QY 247 HNYGHYTIWMDMFGSLRDLPLEDDDK 270
DB 244 HNYGHYTIWMDMFGSLRDLPLEDDDK 267

RESULT 7
09E0S5 PRELIMINARY; PRT; 299 AA.
AC 09E0S5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)

Db 180 YVEVYLPLHLNLEFLVLAFLAVNFWTISIHDOVDFGHHINTTGHHTIHLNCDYGYT 239
 QY 254 IWMWMMFSGSLR-----DPLEEDDKNSDF 277
 Db 240 TWMRLGGSYRPAEOTHOQLSSLLHAGDPKYYDVPYSEYHEKGF 283

RESULT 10

093875 PRELIMINARY; PRT; 386 AA.

AC 093875;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE C5.6 DESATURASE.
 GN ERG3.
 OS Candida albicans (Yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
 OX NCBI_TaxID=5476;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B311;
 RA Miyazaki Y., Bennett J.E., Geber A., Hitchcock C., Parkinson T.,
 RA Miyazaki H.;
 RT "ERG3";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF069752; AAC99343.1;
 DR InterPro: IPR001541; Sterol_desat.
 DR Pfam: PF01598; Sterol_desat; 1.
 SQ SEQUENCE 386 AA; 45447 MW; E5CF1DBB0A12E83B CRC64;

Query Match 19.0%; Score 297; DB 3; Length 386;
 Best Local Similarity 31.4%; Pred. No. 4,4e-18;
 Matches 91; Conservative 46; Mismatches 105; Indels 48; Gaps 14;

QY 10 QPVDSEFYNRIVLSHLPLANLW-EPRLPLOTWLRNYLACTLLYFSGFLMCFYIT--- 65
 Db 97 KFPAPDFIN---TSFLARSKIFRETLSLFLIT---TIFGWLIFYVAIVSYVFPDKK 149
 QY 66 -----VLKINVLPKDAIPITIKARLQMFYAMKAMPYTLTLPVSESMIE-RGWTKCF 118
 Db 150 IFNHPRLK-----NQMSEIKRATTAIVMWLL-TIIPFLLELNKSYSLYL 195
 QY 119 SIDEF--GMILFYVYAIYLVVEEGIYMMHRELHDIKPKYKYLHATGHYINKNQNTLSPE 176
 Db 196 DINCTGYSKALIMQIPKFLITFDGIGYFLHRLH--WPSYKVLHKKPHH--KWICTPF 251
 QY 177 AGLAHFVDDGLQAVPRVIALFTVPIHFTTHIGLLEMEALWTANIHCIHGNITP--PVM 233
 Db 252 ASNAFHFVDDGFOSLPPLHLPPLPLKLYLYLFTFVNFTVMIH--GSYMSNDPVV 308
 QY 234 -GAGYHTIHTTYKKNYGHYITWMDMMFSGSLR--DPLEEDDKNSDFK 279
 Db 309 NGTACHTVHHLHYENYNGOFTTILMDRLGNSRRRDSLSFYKDYAAEEKK 358

RESULT 11
 094457
 ID 094457 PRELIMINARY; PRT; 300 AA.
 AC 094457;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE PUTATIVE C-5 STEROL DESATURASE C1687.16C (EC 1.3.-.-) (STEROL-C5-
 DESATURASE).
 GN SPAC1687.16C.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.

OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RA Wedler H., Wambutt R., Wood V., Rajandream M.A., Barrell B.G.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE INTRODUCTION OF A C-5 DOUBLE BOND IN THE B
 RING OF ERGOSTEROL. MAY CONTRIBUTE TO THE REGULATION OF
 CC ERGOSTEROL BIOSYNTHESIS (BY SIMILARITY).
 CC -1- COFACTOR: IRON (BY SIMILARITY).
 CC -1- PATHWAY: ERGOSTEROL BIOSYNTHESIS.
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC
 CC RETICULUM (BY SIMILARITY).
 CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
 CC AND/OR BE INVOLVED IN METAL ION BINDING.
 CC -1- SIMILARITY: TO OTHER FATTY ACID DESATURASES.
 DR EMBL: AL035064; CAA22610.1;
 DR InterPro: IPR001541; Sterol_desat.
 DR Pfam: PF01598; Sterol_desat; 1.
 KW Hypothetical protein; Sterol biosynthesis; Oxidoreductase;
 KW Endoplasmic reticulum; Iron; Transmembrane.
 FT TRANSMEM 33 53
 FT TRANSMEM 117 137
 FT TRANSMEM 180 200
 FT DOMAIN 137 142
 FT DOMAIN 150 154
 FT DOMAIN 224 229
 FT DOMAIN 224 229
 SQ SEQUENCE 300 AA; 35625 MW; 8F0E08972CE29EF9 CRC64;

Query Match 17.9%; Score 279.5; DB 3; Length 300;
 Best Local Similarity 27.6%; Pred. No. 1.1e-16;
 Matches 76; Conservative 53; Mismatches 107; Indels 39; Gaps 11;

QY 7 YLMQVDE--TSFYNRIVLSHLPLANLW-EPRLPLOTWLRNYLACTLLYFSGFLMCF 62
 Db 3 YLNAVDAQYALDSIYNVY--PLARDNIVROSISLFLTW---FGMFF.YLTFASLSYQ 56
 QY 63 YIY-----YKINVLPKDAIPITIKARLQMFYAMKAMPYTLTLPVSESMIE-RG 112
 Db 57 FVFDKSLMDHPRKPLKNQV-----MEVLATLQNLPMALL-VPMFLAEHLG 102
 QY 113 WTKFASIDEGWILFYVYAIYLVVEEGIYMMHRELHDIKPKYKYLHATGHYINKNQNT 172
 Db 103 YSLYDINISDGLKYLFCSLPLPFMSDFGIYMAHRELH--RIVYRPLHLH--KWII 158
 QY 173 LSPFAGLAFHPVDDGLQAVPRVIALFTVPIHFTTHIGLLEMEALWTANIHCIHGNITPV 232
 Db 159 CTYASASHAFKSDGFLQSLPPLHLPPLPPLKLYLYLFTFVNFTVMIHIDGKYSNNPI 218
 QY 233 M-GAGYHTIHTTYKKNYGHYITWMDMMFSGSLRDP 266
 Db 219 INGAHHNGHHIYENYNGOFTTILDRLGNSFRAP 253

RESULT 12
 09LNMW2
 ID 09LNMW2 PRELIMINARY; PRT; 261 AA.
 AC 09LNMW2;
 DT 01-OCT-2000 (TReMBLrel. 15, Created)
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
 DE P22G5.23.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eud.; cots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,
 RA Kim C., Altafi H., Bel O., Chin C., Chlou J., Choi E., Conn L.,
 RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,

RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F2265 from chromosome
RT 1";
RN Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RL
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RN Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RL
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RN Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL
RP SEQUENCE FROM N.A.
RA Cheuk R., Shin P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E.,
RA Com L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RN Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RL
DR EMBL: AC022464; AAF79571.1; -;
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
SQ SEQUENCE 261 AA; 30493 MW; 83DBE0A795CC749F CRC64;

Query Match 11.3%; Score 176; DB 10; Length 261;
Best Local Similarity 22.6%; Pred. No. 9.8e-08;
Matches 65; Conservative 33; Mismatches 77; Indels 112; Gaps 14;
QY 42 WLRNYLA-----GTL-----YFISGFLMCF----- 62
DB 2 WLMQVLYHPSDQLACISGFLHESVFLSGLPFTLEKOGLSKKYKTKNNTPAAG 61
QY 63 -----YIYUUKINVLPRDAIPTIKAMRLQMFVAMKAMFWYTLPTVSESMTERGWTKC 116
DB 62 KCTIRLLVHFVSVPRLMLASVPRAMGMR-----SSPF-----LPS-----WKVEY 103
QY 117 FASIDFGMLVGVYTAIVLVEEGITWMMHRELHDIKPLKYLKATLNTIYNKONTLSPP 176
DB 104 SAQ-----ILFFIILEDVFWGCHRLHS-KWLYKNVSHVNHYY-----ATPF 145
QY 177 A-----GLAFHRYDGLQAVPHYIALFIVPHHTTHIGLFPMAITANIHDC 224
DB 146 GLTSEVANHRAELFLGFATVIGPALTGPHLITLWL-----WVLRVLETVENHCG 195
QY 225 IH-----GNIMPYM-GAGYHTTHT--TYKHNVGHTYIMDMFGS 262
DB 196 YHRPWSLNSNRLPLYGADPHDYNHRLLYTKSGNYSFVYMDMIFCG 242

RESULT 13
Q20612 PRELIMINARY; PRT; 286 AA.
AC Q20612;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE F49E12.10 PROTEIN.
GN F49E12.10.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Felodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Thomas K.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berke M.,
RA Bonfield J., Burton J., Connell M., Copsy T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Klirsten J., Laister N., Latreille P.,
RA Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smaildon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstock L., Wilkinson-Sproat J., Wohlman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans.";
RL Nature 368:32-38(1994).
DR EMBL: Z66520; CAA91384.1; -;
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
SQ SEQUENCE 286 AA; 33874 MW; EA9A7468A98D8817 CRC64;

Query Match 11.2%; Score 174.5; DB 5; Length 286;
Best Local Similarity 22.4%; Pred. No. 1.5e-07;
Matches 64; Conservative 49; Mismatches 112; Indels 61; Gaps 14;
QY 39 LQTVLRNYLAGT--LLYFTSG-----FLWCFYIYUUKINVLPRDAIPTIKAMRLQMFV 90
DB 5 LYRLKYNFEGDGYMLYVGVGNAGISFMYLNLFFIIIDVDPK----WVQYKQDEK 60
QY 91 AMKAMPWYTLPTVSES-----IERGTFKCPASIDEGGMLYVYAIYV 137
DB 61 KPSLSYULSLTKVGNOLIVPIVITLWYVARMGMPGVPYIP-SWYILRLDACLMA 119
QY 138 FVEEGITWMMHRELHDIKPLKYUHLATHNIYNKONTSPFAGLAFHDPDGLQAVPHYAL 197
DB 120 MEDIGFYTHRLFNPK-LYKHKKHKNEMAPVSI--SIANLE--NAISLSPI 172
QY 198 FIVPHHTTHI--GLLFEAIVTANIHDCIGNIMPYM-GAGYHTTHTIYKNYGHYT 253
DB 173 YLCAVLEFRCHVSHYLFYSVALITTFEHN--SGYHPRFMLSABHNDHFNKVFNECYGCFP 230
QY 254 IYMDMFG-----SLRDP-----LLEEDNKDSFK 279
DB 231 --LDWLHGDTFRKSIENHPRDYVYVYGTPIKELIPEDKNNNNKK 274

RESULT 14
Q90010 PRELIMINARY; PRT; 115 AA.
AC Q90010;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE HYPOTHETICAL PROTEIN (FRAGMENT).
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=968 H90.
RA Ding D.;
RT "Generation and analysis of GFP-gene fusion library of fission
RT yeast.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB027881; BAA87185.1; -;
DR InterPro: IPR001541; Sterol_desat.
FT NON_TER 1 115
FT NON_TER 1 115
SQ SEQUENCE 115 AA; 13616 MW; F8E7A05958C23D23 CRC64;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 17:06:27 ; Search time 9.86 Seconds

(without alignments)
1044.910 Million cell updates/sec

Title: US-09-775-879-21

Perfect score: 1562
Sequence: 1 MADNAYLMQFVDETSFYNR.....SLRDPLEEDNDKDSFKAE 281

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt.39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	341	21.8	299	1 SC5D_MOUSE	088822 mus musculus
2	330	21.1	299	1 SC5D_HUMAN	075845 homo sapien
3	309.5	19.8	364	1 ERG3_CANGA	P50860 candida gla
4	304.5	19.5	329	1 ERG3_SCHPO	O13666 schizosacch
5	295.5	18.9	365	1 ERG3_YEAST	P32353 saccharomyc
6	156	10.0	293	1 ER25_HUMAN	O15800 homo sapien
7	152.5	9.8	309	1 ER25_YEAST	P53045 saccharomyc
8	141.5	9.1	308	1 ER25_CANAL	O59933 candida alb
9	117	7.5	349	1 SUR2_YEAST	P38992 saccharomyc
10	98	6.3	615	1 YCH0_YEAST	P25956 saccharomyc
11	98	6.3	615	1 YK86_YEAST	P36173 saccharomyc
12	92.5	5.9	542	1 LMT_YEAST	O90188 chlamydia m
13	91	5.8	534	1 TNAL_YEAST	P53222 saccharomyc
14	90.5	5.8	279	1 YR6_MYCCA	P43045 mycoplasma
15	90	5.8	265	1 COX3_MARPO	P26858 marchantia
16	89	5.7	649	1 COX1_BACSU	P16365 bacillus su
17	88.5	5.7	308	1 COX3_MOUSE	O21403 struthio ca
18	87.5	5.6	261	1 COX3_STRCA	P92896 pongo pygma
19	87	5.6	261	1 COX3_PONRA	P33303 anopheles q
20	87	5.6	341	1 CYB_GYMTI	P29637 gymnotina
21	86	5.5	308	1 CYB_ANOUU	O76024 homo sapien
22	86	5.5	890	1 WFS1_HUMAN	O08675 mus musculu
23	85	5.4	369	1 PAR3_MOUSE	O84539 chlamydia t
24	85	5.4	542	1 LMT_CHLTR	O37620 protoliteca
25	84.5	5.4	262	1 COX3_PROMT	O09555 caenorhabdi
26	84.5	5.4	805	1 DF19_CAEEL	P80441 rhizopus st
27	84	5.4	281	1 COX3_RHIST	P92514 arabidopsis
28	83.5	5.3	265	1 COX3_ARATH	O03215 chromallium
29	83.5	5.3	411	1 CYB_CHRVI	O07285 saccharomyc
30	83.5	5.3	1511	1 PDRC_YEAST	P51751 rhodospirill
31	83	5.3	255	1 RCBM_RHOPH	P28476 homo sapien
32	83	5.3	465	1 GAR2_HUMAN	P25628 saccharomyc
33	83	5.3	610	1 ARE1_YEAST	

34	82.5	5.3	353	1 CKR8_MOUSE	P56484 mus musculus
35	82.5	5.3	580	1 P69_MYCHR	P15362 mycoplasma
36	82	5.2	338	1 YVAD_BACSU	P37520 bacillus su
37	81.5	5.2	329	1 COX3_ACACA	O37374 acanthamoeb
38	81.5	5.2	581	1 FUR4_SCHPO	O10279 schizosacch
39	81.5	5.2	2273	1 HPA1_YEAST	P32874 saccharomyc
40	81	5.2	269	1 COX3_EMENT	P00421 emeticella
41	81	5.2	560	1 YG14_YEAST	P53154 saccharomyc
42	81	5.2	824	1 SIIP_SALTY	O92HC7 salmonella
43	80.5	5.2	374	1 PAR3_HUMAN	O00254 homo sapien
44	80.5	5.2	392	1 CYB_SOLTU	P29757 solanum tub
45	80.5	5.2	465	1 GAR2_MOUSE	P56476 mus musculus

ALIGNMENTS

RESULT ID	SC5D_MOUSE	STANDARD	PRT	299 AA.
AC	088822:			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	20-AUG-2001 (Rel. 40, Last annotation update)			
DE	LATHOSTEROL OXIDASE (EC 1.3.3.2) (LATHOSTEROL 5-DESATURASE) (STEROL-5-DESATURASE)			
GN	SC5DL OR SC5D			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;			
OX	NCBI_Taxid=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-BALB/C;			
RA	Nishi S., Nishino H., Ishibashi T.;			
RT	"Molecular cloning and expression of the human and mouse lathosterol 5-desaturase."			
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.			
CC	- 1- CATALYTIC ACTIVITY: 5-ALPHA-CHOLEST-7-EN-3-BETA-OL + O(2) =			
CC	CHOLESTA-5,7-DIEN-3-BETA-OL + H(2)O(2).			
CC	- 1- COFACTOR: IRON (BY SIMILARITY).			
CC	- 1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC			
CC	RETICULUM (PROBABLE).			
CC	- 1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE			
CC	AND/OR BE INVOLVED IN METAL ION BINDING.			
CC	- 1- SIMILARITY: BELONGS TO THE STEROL DESATURASE FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL: AB016248; BAA33730.1; -			
DR	MGD: MGI:1353611; SC5d.			
DR	InterPro: IPR001541; Sterol_desat.			
DR	Plan: PFO1598; Sterol_desat; 1.			
KW	Sterol biosynthesis; Oxidoreductase; Endoplasmic reticulum; Iron;			
KW	Transmembrane.			
FT	TRANSMEM 32 52			POTENTIAL.
FT	TRANSMEM 79 99			POTENTIAL.
FT	TRANSMEM 117 137			POTENTIAL.
FT	TRANSMEM 186 206			POTENTIAL.
FT	DOMAIN 138 143			HISTIDINE BOX 1.
FT	DOMAIN 151 155			HISTIDINE BOX 2.
FT	DOMAIN 228 233			HISTIDINE BOX 3.
SO	SEQUENCE 299 AA; 35062 MW; C6B3BBAA44D58E177 CRC64;			

Query Match 21.8%; Score 341; DB 1; Length 299;
Best Local Similarity 33.2%; Pred. No. 1.4e-19;

[illegible]

```

RESULT 2
CC SC5D_HUMAN STANDARD: PRT; 299 AA.
AC 075845; 000119;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE LATHOSTEROL OXIDASE (EC 1.3.3.2) (LATHOSTEROL 5-DESATURASE) (DELTA-7-
DE STEROL 5-DESATURASE) (C-5 STEROL DESATURASE) (STEROL-C5-DESATURASE).
GN SC5DL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Nishi S., Nishino H., Ishibashi T.;
RT "Molecular cloning and expression of the human and mouse lathosterol
RT 5-desaturase.";
RL Submitted (JUL-1998) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97130614; Pubmed=8976377;
RA Matsushima M., Inazawa J., Takahashi E., Suzumori K., Nakamura Y.;
RT "Molecular cloning and mapping of a human cDNA(SC5DL) encoding a
RT protein homologous to fungal sterol-C5-desaturase.";
RL Cyogenet. Cell Genet. 74:252-254(1996).
CC -1- CATALYTIC ACTIVITY: 5-ALPHA-CHOLEST-7-EN-3-BETA-OL + O(2) =
CC -1- CHOLESTA-5,7-DIEN-3-BETA-OL + H(2)O(2).
CC -1- COFACTOR: IRON (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM (PROBABLE).
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/OR BE INVOLVED IN METAL ION BINDING.
CC -1- SIMILARITY: BELONGS TO THE STEROL DESATURASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AB016247; BAA33729.1; -
DR EMBL: DB5181; BAA18970.1; -.
DR MIM: 602286; -
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.

```

Accession	Protein	Length (aa)	Weight (kDa)	PI	Source
KM	Sterol biosynthesis; Oxidoreductase; Endoplasmic reticulum; Iron;	299	35.261	5.2	POTENTIAL.
KM	Transmembrane.	32		5.2	POTENTIAL.
FT	TRANSMEM	79		9.9	POTENTIAL.
FT	TRANSMEM	117		13.7	POTENTIAL.
FT	TRANSMEM	186		20.6	POTENTIAL.
FT	DOMAIN	138		14.3	HISTIDINE BOX 1.
FT	DOMAIN	151		15.5	HISTIDINE BOX 2.
FT	DOMAIN	228		23.3	HISTIDINE BOX 3.
FT	CONFLICT	216		29.9	POLYDOPINGSHHTDHDHMFDPYNGOYFTLMDIRGGSFKN PSSEFGKPLSVKMEFGKRSRSPGCKNCKNKLKNGEFTK TE-> RAKNYSMEQLDRLNRLPSYS (IN REF. 2).
SQ	SEQUENCE	299	AA:	35261	MM: 9EF8B21EE3EFAA56 CRC64;

Query Match	21.1%	Score 330;	DB 1;	Length 299;
Best Local Similarity	32.9%	Pred. No. 1e-18;		
Matches 82; Conservative	52;	Mismatches 73;	Indels 42;	Gaps 12;

```

QY 51 LLYEYSGEFLMCYYIV-----YKLIWVYPKDAIPFKAMRIOMFAAMKAMPYTL 101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 44 ILYFCATLSYEVFDHMLMKHPQFLKNQV-----RREIKFTVQALPMISL 90
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 102 PLYSESMTI-RCWTKCFASIDEEGMLYRVYTAI--YLVFEFGYVMHRELDIKPLXK 158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 91 -TVALELELEIRYSKLDHDLGEFPGLEFLVYIISLFLEFTDMFLYIMHRLGHN-RLVYK 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 155 YLHAHHHTYKRNKNTLSPFAGLAFHHYVDGLQAVPHVIALEYFPIHFTTHIGLLFMEALWT 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 149 RLKHKRHNI--KIPFPASHAFHPIDELQSLPHIHYPIELPLKVVYLSLYLIVLWIT 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 219 ANIHDCIHGN-----IMD-VMGACHTHTHTTYKHNGHY-TIMDMWFGSLRDLLEE 270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 206 TSIH--GDFEVPOILOPFIINGSAAHHTDHNMEFDYNYGQYETLW-DRIGSEFKNDSFE 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 DDNKDSFKK 279
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 CKGPLSYVK 270
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 3
ERG3_CANGA STANDARD: PRT: 364 AA.
ID ERG3_CANGA
AC P50860:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE C-5 STEROL DESATURASE (EC 1.3.-.) (STEROL-C5-DESATURASE).
GN ERG3.
CS Candida glabrata (Yeast) (Torulopsis glabrata).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_taxid=5478;
[1]
RP SEQUENCE FROM N.A..
RC STRAIN-2001-15.
RX MEDLINE=96161286; Pubmed=8593007;
RA Geber A., Hitchcock C.A., Swartz J.E., Pullen F.S., Marsden K.E.,
RA Kwon-Chung K.J., Bennett J.E.;
RT "Deletion of the Candida glabrata ERG3 and ERGL genes: effect on cell
RT viability, cell growth, sterol composition, and antifungal
RT susceptibility."
RL Antimicrob. Agents Chemother. 39:2708-2717(1995).
CC -1- FUNCTION: CATALYZES THE INTRODUCTION OF A C-5 DOUBLE BOND IN THE B
CC RING OF ERGOSTEROL. MAY CONTRIBUTE TO THE REGULATION OF
CC ERGOSTEROL BIOSYNTHESIS.
CC -1- COFACTOR: IRON (BY SIMILARITY).
CC -1- PATHWAY: ERGOSTEROL BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM (PROBABLE).
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/OR BE INVOLVED IN METAL ION BINDING.
CC -1- SIMILARITY: BELONGS TO THE STEROL DESATURASE FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L40390; AAB02330.1; -
DR InterPro; IPR001541; Sterol_desat.
DR Pfam; PF01598; Sterol_desat; 1.
KM Sterol biosynthesis; Oxidoreductase; Endoplasmic reticulum; Iron;
KM Transmembrane.
FT TRANSMEM 94 114 POTENTIAL.
FT TRANSMEM 142 162 POTENTIAL.
FT TRANSMEM 249 269 POTENTIAL.
FT DOMAIN 201 205 HISTIDINE BOX 1.
FT DOMAIN 214 218 HISTIDINE BOX 2.
FT DOMAIN 289 293 HISTIDINE BOX 3.
SQ SEQUENCE 364 AA; 42735 MW; 91010FD86BC137F0 CRC64;

Query Match 19.8%; Score 309.5; DB 1; Length 364;
Best Local Similarity 31.2%; Pred. No. 4.6e-17;
Matches 89; Conservative 46; Mismatches 103; Indels 47; Gaps 11;

OY 7 YLMQVDETSFYNRIVLSHLPLANMEPLPHLOTWLRNYLAGTLTYFSGFLWCYIY- 65
DB 72 FLEDETEKT-----YASLLEPRN--NLIREFESLAAVTVGLLLITASISYVEVD 122
OY 66 -----YLKINVYLPKRAIPTIKAMRLQMFVAMKAMPWYLLPTVSSSME-RGWTGC 116
DB 123 RIFNHPKILKQMW-----LEKLAVSATPTMSLL-TVPMFLMLNCYISKI 168
OY 117 FASID--EFGWILYFVYIAYLVFEVEGTYMMHRELHDIKPLKYLAATHHYNKONTLS 174
DB 169 YVDVDMENHGLKRLILEVATFTFFTCGIYLAHRMLWMPR-VYKALHKPHN--KMLVGC 224
OY 175 PRAGLAFFHYVDGILQAVPHVIALFTVPIHFTTHIGLLEMAIWTANIHCINIPVM- 233
DB 225 PRASHAFHFVDYFQSLASHIYPMILPLKHSYLLIFTEVFNWSVMIHGQHSNNPVNA 284
OY 234 GAGYFTIHHTTKHNYGHTYIMDMWFGSLR-----DPLEED 271
DB 285 GTACHTVHHLFNYNYGFTLTMDRIGSYRPPEDSLFPKLMAD 329

RESULT 4
ERG3_SCHPO STANDARD; PRT: 329 AA.
AC 013666;
DT 15-JUL-1999 (Rel. 38; Created)
DT 15-JUL-1999 (Rel. 38; Last sequence update)
DT 20-AUG-2001 (Rel. 40; Last annotation update)
DE PROBABLE C-5 STEROL DESATURASE (EC 1.3.-.-) (STEROL-C5-DESATURASE).
GN SPC27812.03C OR P1075.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Kushiida N., Yamazaki S., Tanaka T., Jinno K., Haikawa Y., Yamazaki J.,
RA Yamamoto S., Sekine M., Oguchi Y., Nagai Y., Sakai M., Aoki K.,
RA Ogura K., Otsuka R., Kudoh Y., Yanagida M., Machida M., Zhang M.Q.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Rajadream M.A., Barrell B.G., Lauber J., Hilbert H.,

```

```

RA Duesterhoeft A.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: CATALYZES THE INTRODUCTION OF A C-5 DOUBLE BOND IN THE B
CC RING OF ERGOSTEROL. MAY CONTRIBUTE TO THE REGULATION OF
CC ERGOSTEROL BIOSYNTHESIS (BY SIMILARITY).
CC - COFACTOR: IRON (BY SIMILARITY).
CC - PATHWAY: ERGOSTEROL BIOSYNTHESIS.
CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM (PROBABLE).
CC - DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/OR BE INVOLVED IN METAL ION BINDING.
CC - SIMILARITY: BELONGS TO THE STEROL DESATURASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB004539; BAA21457.1; -
DR EMBL; AL021766; CAA16898.1; -
DR InterPro; IPR001541; Sterol_desat.
DR Pfam; PF01598; Sterol_desat; 1.
KM Sterol biosynthesis; Oxidoreductase; Endoplasmic reticulum; Iron;
KM Transmembrane.
FT TRANSMEM 67 87 POTENTIAL.
FT TRANSMEM 149 169 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
FT DOMAIN 170 175 HISTIDINE BOX 1.
FT DOMAIN 183 187 HISTIDINE BOX 2.
FT DOMAIN 257 262 HISTIDINE BOX 3.
SQ SEQUENCE 329 AA; 39063 MW; C9DEF93732BC4BFA CRC64;

Query Match 19.5%; Score 304.5; DB 1; Length 329;
Best Local Similarity 31.2%; Pred. No. 1e-16;
Matches 85; Conservative 53; Mismatches 107; Indels 27; Gaps 12;

OY 3 ADNAYLMQVDETSFYNRIVLSHLPLA-NLMEPLPH-FLQTLWRNYLAGTLTYFSGFLW 60
DB 34 AVNSTGLGLEAKVNF--AITSGLLDNRNNVWRQTSFLITW---LMTLSYFLSA-SF 85
OY 61 CFYIYIYIKINIVLPKRAIPTIKAMR-----LQMFVAMKAMPWYLLPTVSSSME-RGWTGC 115
DB 86 AYVYFDR-----EEARRHPRKFLKQNHLELWALKNLPGMAIL--TAPPELAEIRNGY 138
OY 116 CFASIDFEGWILYFVYIAYLVFEVEGTYMMHRELHDIKPLKYLAATHHYNKONTLS 175
DB 139 VYKLDYEGYFYLFPSIALDLFSDFLITWIRALNH-RWLYAPLHLNH---KWIVPTP 194
OY 176 FAGLAFFHYVDGILQAVPHVIALFTVPIHFTTHIGLLEMAIWTANIHCINIPVM- 234
DB 195 YSSHAFFHYLDYGQSPLPYHMFPEFFPLNKYVYLLPESQVWYVYLIRDGAFYSNNAVNG 254
OY 235 AGYHTIHHTTKHNYGHTYIMDMWFGSLR 266
DB 255 AAHHAAHHMFNYNYGFTLFPRLCSSYROP 286

RESULT 5
ERG3_YEAST STANDARD; PRT: 365 AA.
AC P32353;
DT 01-OCT-1993 (Rel. 27; Created)
DT 01-OCT-1993 (Rel. 27; Last sequence update)
DT 20-AUG-2001 (Rel. 40; Last annotation update)
DE C-5 STEROL DESATURASE (EC 1.3.-.-) (STEROL-C5-DESATURASE).
GN ERG3 OR STR1 OR YLR056W OR L2150.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetaceae.

```

```

OX NCBI_TaxID=4932;
RX [1] SEQUENCE FROM N.A.
RA MEDLINE=91323727; PubMed=1864507;
RA Archington B.A., Bennett L.G., Skatrud P.L., Guynn C.J., Barduch R.J.,
RA Ulbright C.E., Bard M.;
RA "Cloning, disruption and sequence of the gene encoding yeast C-5
RA sterol desaturase";
RA Gene 102:39-44(1991).
RN [2]
RN SEQUENCE FROM N.A.
RA Renaud G., Lacroite F., Cassart J.P., Vandenhoute J., Delcour J.;
RA Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
RL [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=94236239; PubMed=8180700;
RA Taguchi N., Takano Y., Julmanop C., Wang Y., Stock S.,
RA Takemoto J.Y., Miyakawa T.;
RA "Identification and analysis of the Saccharomyces cerevisiae SVR1
RA gene reveals that ergosterol is involved in the action of
RA syringomycin.";
RL Microbiology 140:353-359(1994).
RN [4]
RP SEQUENCE FROM N.A.
RA Andre B., Urrestarazu L.A.;
RA Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.
RL -1- FUNCTION: CATALYZES THE INTRODUCTION OF A C-5 DOUBLE BOND IN THE B
CC RING OF ERGOSTEROL. MAY CONTRIBUTE TO THE REGULATION OF
CC ERGOSTEROL BIOSYNTHESIS.
CC -1- COFACTOR: IRON (BY SIMILARITY).
CC -1- PATHWAY: ERGOSTEROL BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. ENDOPLASMIC
CC RETICULUM (PROBABLE).
CC -1- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/OR BE INVOLVED IN METAL ION BINDING.
CC -1- SIMILARITY: BELONGS TO THE STEROL DESATURASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M62623; AAA34594.1; -
DR EMBL; S46162; AAB39844.1; -
DR EMBL; M64989; AAA34595.1; -
DR EMBL; D14299; BAA20292.1; -
DR EMBL; X94607; CAA64303.1; -
DR EMBL; Z73228; CAA97586.1; -
DR PIR; J01146; J01146.
DR PIR; S41993; S41993.
DR SGD; S0004046; ERG3.
DR InterPro; IPRO01541; Sterol_desat.
DR Pfam; PF01598; Sterol_desat; 1.
KW Sterol biosynthesis; Oxidoreductase; Endoplasmic reticulum; Iron;
KW Transmembrane.
KM TRANSMEM 93 113
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 243 263 POTENTIAL.
FT DOMAIN 200 204 HISTIDINE BOX 1.
FT DOMAIN 213 217 HISTIDINE BOX 2.
FT DOMAIN 288 292 HISTIDINE BOX 3.
SQ SEQUENCE 365 AA; 42730 MW; 7F441DA6927A711C CRC64;

```

```

Query Match 18.9%; Score 295.5; DB 1; Length 365;
Best Local Similarity 31.5%; Pred. No. 5.5e-16;
Matches 87; Conservative 43; Mismatches 95; Indels 51; Gaps 12;
15 TSPFNRIIVLSHLPANMEPLPHFLQITWLRNYLAGTLLYFISGFIMCRFYI----- 65
||||| : ||| : | || : | ||| : | : :

```

```

DB 78 TSP-----ASLPRS--SILREFLSLWIVITIEGLLLYFTASLSYVEVFDKSIINHP 129
OY 66 YLKINVLPRDAIPTKARLQMFAMKAMPWTLPTVSESME-RCGTKCFASIDERG 124
||| : : : : : : : : : : : : : : : : :
DB 130 YLK-----NQAMEIKLAVSAIPMMSML-TYPWPFMELNGSKLYMKIDYEN 175
OY 125 W-----ILFEYVIAIYLVFEFGIYWMHRELHDIKPLKYLAHNIYKNTLSPFAG 178
||| : : : : : : : : : : : : : : : : :
DB 176 HGVRKLLIEYFTFI-----FFTDGCVYLAHRLWMPR-VYRALHKPHH---KWLVCYFPAS 227
OY 179 LAFHPVDGILQAVPHVIALFPIVPIHTTHIGLFLMEAIWTANIHDCIH-GNIMPVAGY 237
||||| : : : : : : : : : : : : : : : : :
DB 228 HSFHPVDGLOSTSYHIIYLLPLKAVSYLLFTFVNFPTWMTIDGQYLSNPNVAVGTAC 287
OY 238 HTIHTTYKKNYGHVYTIWDMWFGSLRDLPLEDDN 273
||| : : : : : : : : : : : : : : : : :
DB 288 HTVHILYFNVNYGQFTTLMDRLGGSYRRP-----DGS 319
RESULT 6
ER25_HUMAN STANDARD: PRT; 293 AA.
AC Q15800;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE C-4 METHYL STEROL OXIDASE (EC 1.1.1.11).
GN SC4MOL OR ERG25 OR DESP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96279274; PubMed=8663358;
RA LI L., Kaplan J.;
RA "Characterization of yeast methyl sterol oxidase (ERG25) and
RA identification of a human homologue.";
RN J. Biol. Chem. 271:16927-16933(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RA Herrmann K.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- COFACTOR: IRON (PROBABLE).
CC -1- PATHWAY: CHOLESTEROL BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-ASSOCIATED; ENDOPLASMIC RETICULUM.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; U60205; AAC50587.1; -
DR EMBL; U93162; AAB81566.1; -
DR InterPro; IPRO01541; Sterol_desat.
DR Pfam; PF01598; Sterol_desat; 1.
KW Sterol biosynthesis; Oxidoreductase; Transmembrane; Iron.
KM CARBOHYD 4 4 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 199 219 POTENTIAL.
SQ SEQUENCE 293 AA; 35215 MW; D86EDDBE85DE0BF CRC64;

```

```

Query Match 10.0%; Score 156; DB 1; Length 293;
Best Local Similarity 22.8%; Pred. No. 2.4e-05;
Matches 74; Conservative 41; Mismatches 110; Indels 100; Gaps 16;

```



```

QY 9 MOFVDETFSEYNRIVLSHLLPAN-LMEPLPHLOTWLRNYLA-----GTL-----LYF-- 54
DB 17 VEYVDS-----LLPENPLDPEFKMANMYLNNTTKQIATWGSLLIHEALYELF 65
QY 55 -ISGFLKCFYIYKINLYLPKDAIPT-----IKARLOMF----- 89
DB 66 CLPGLFQIFIPKMKYKI--OKDKPEYEMOMKCKFVLLFNHFCIOLPLICGTYPTETEF 123
QY 90 ---VAMKAMP--WYLLPVSSESMIERGWTGCF--ASIDEGWLLFYVYAIYLVFVEFGIY 144
DB 124 NIPYWEKMPRYFLL-----ARCGCAVIEDTW-----HY 154
QY 145 WMHRELHDIKPLKYLAHTHHIYNKONTLSPFAGLA--FHPYDGILOAVPHYIALEFVPI 202
DB 155 FLHRLHH-KRIYKYIHKVHNEFQ-----APFGMEAEVHPILETLLIGFFIIGIYLLCD 208
QY 203 HFTTHIGLFMEAIWTANIHDCIHGNIMPV-----MGAGYPTIHHTTYKHNHYGHTYIM 256
DB 209 HVILLMAVVTIRLLETIVHSGYDIPLNPLNLIPYAGSRHHDHFMNFIQVASTFTWM 268
QY 257 DMHGGSLDPLLEDDNKDSFKKAE 281
DB 269 DRIFGTSQYANAYNKKRKKFEKTE 293

RESULT 7
ER25_YEAST STANDARD: PRT: 309 AA.
AC P53045;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE C-4 METHYL STEROL OXIDASE (EC 1.-.-.-).
GN ERG25 OR FET6 OR YGR060W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96133902; PubMed=8552601;
RA Bard M., Bruner D.A., Pierson C.A., Lees N.D., Biermann B., Frye L.,
RA Koegel C., Barbuch R.;
RT "Cloning and characterization of ERG25, the Saccharomyces cerevisiae
RT gene encoding C-4 sterol methyl oxidase.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:186-190(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Entian K.D., Rose M., Koetter P., Roehmer A., Sehmsam I., Hempel S.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX MEDLINE=96279274; PubMed=8663358;
RA Li L., Kaplan J.;
RT "Characterization of yeast methyl sterol oxidase (ERG25) and
RT identification of a human homologue.";
RL J. Biol. Chem. 271:16927-16933(1996).
CC -I- FUNCTION: CATALYZES THE FIRST STEP IN THE REMOVAL OF THE TWO C-4
CC METHYL GROUPS OF 4,4-DIMETHYLAZYMOSTEROL.
CC -I- COFACTOR: IRON.
CC -I- PATHWAY: ERGOSTEROL BIOSYNTHESIS.
CC -I- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM (PROBABLE).
CC -I- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/OR BE INVOLVED IN METAL ION BINDING.
CC -I- SIMILARITY: BELONGS TO THE STEROL DESATURASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).

```

```

CC -----
DR EMBL: U31885; AAC49139.1;
DR EMBL: 272845; CAA97062.1;
DR SGD: S0003292; ERG25.
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
KW Sterol biosynthesis; Oxidoreductase; Endoplasmic reticulum;
KW Transmembrane; Iron.
FT TRANSMEM 56 76
FT DOMAIN 160 164
FT DOMAIN 173 177
FT DOMAIN 257 263
SQ SEQUENCE 309 AA; 36479 MW; E0422CA16DD17794 CRC64;

Query Match 9.8%; Score 152.5; DB 1; Length 309;
Best Local Similarity 20.4%; Pred. No. 4,6e-05;
Matches 68; Conservative 55; Mismatches 111; Indels 99; Gaps 16;

QY 15 TSFYNRIVLSHLLPANL-----WEPLPHLO-----TWLRNYAGTLVFISS 57
DB 2 SAVFNNAITSLGLVQASTYSQTLONVAHYOPOLNFMERYMAWYSYNNNDVATGLMF--- 58
QY 58 FLWCFTIYLYLKINLYLPKDAIPTIKARLOMFVAMKAMPYLLPVSSESMIERGWTGCF 117
DB 59 FLHHEFMYFRCPLPWFIIIDIPYFRKKLOP-----TKIPSAKKDLY-----CL 102
QY 118 AST-----DEFGWLLFYVYAIYLVFVERGI-----YMHRE 149
DB 103 KSVLLSHFLVEAIPITWTFHPCERKLTGIVEVPPPSLKTMALEIGLFVLBDTWYMAHRL 162
QY 150 LHDIKRLKYLAHTHHIYNKONTLSPFAGLA--FHPYDGI-----LOAVPHYIALEFV 201
DB 163 FH-YGVFYKRIHKQHRV-----AAPG-GSAAYAHAEFLSLGFGTVSGPILYVMTGK 215
QY 202 IHFTT---HIGLFMEAI-----WTANIHDCIHGNIMPV-MGAGYPTIHHTTYKHN 248
DB 216 LHFILCVMTLRLPQAVDSHSGYDFPWSLN-----KIMPMWAGAEHHHDHNFYIGN 268
QY 249 YGHYTTWMDMGGSLDPLLEDDNKDSFKKAE 281
DB 269 YASSFRWDYCLDTESGPEAKASREBRMKRAE 301

RESULT 8
ER25_CANAL STANDARD: PRT: 308 AA.
AC O59933;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE C-4 METHYL STEROL OXIDASE (EC 1.-.-.-).
GN ERG25.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=CAI-8;
RA Johnson T., Kennedy M.A., Lees N.D., Bard M.;
RT "ERG25, C-4 sterol methyl oxidase from Candida albicans.";
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: CATALYZES THE FIRST STEP IN THE REMOVAL OF THE TWO C-4
CC METHYL GROUPS OF 4,4-DIMETHYLAZYMOSTEROL (BY SIMILARITY).
CC -I- COFACTOR: IRON (BY SIMILARITY).
CC -I- PATHWAY: ERGOSTEROL BIOSYNTHESIS.
CC -I- SUBCELLULAR LOCATION: ENDOPLASMIC RETICULUM (PROBABLE).
CC -I- DOMAIN: THE HISTIDINE BOX DOMAINS MAY CONTAIN THE ACTIVE SITE
CC AND/OR BE INVOLVED IN METAL ION BINDING.
CC -I- SIMILARITY: BELONGS TO THE STEROL DESATURASE FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF051914; AAC06014.1; -
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
KW Sterol biosynthesis; Oxidoreductase; Endoplasmic reticulum;
KW Transmembrane; Iron.
FT TRANSMEM 56 76 POTENTIAL.
FT DOMAIN 160 164 HISTIDINE BOX 1.
FT FT 173 177 HISTIDINE BOX 2.
FT DOMAIN 257 263 HISTIDINE BOX 3.
SQ SEQUENCE 308 AA; 36561 MW; 45D7D7AE4081BC15 CRC64;

Query Match 9.18; Score 141.5; DB 1; Length 308;
Best Local Similarity 22.3%; Pred. No. 0.00032;
Matches 65; Conservative 38; Mismatches 95; Indels 93; Gaps 18;

QY 59 LMCFEYIYLKINYLPRDAIPTIKAMRLQMEVAMKAMPWYTL-----LPT 103
DB 39 LKSGYIYMANDLF-----ATGLFELTHEIEFGRCLEPMALIDRIPYFRKKKIDDEKIPS 94
QY 104 VSESMIERGWTGCFAS-----IDERG-WLY-----FVYIAIY 135
DB 95 DKEDQ-----W-ECLKSVLTLSHFLVEAFRPFMFHPLCKIGISYGVPPPKITDMLQNAVF 148
QY 136 LVVEFGIYMKHRELNDIKLYKIATHTIYNKONTLSPAGAA--FHVDCIL----- 188
DB 149 FVLEDIWMYHFNHGLH-YGVFYKXIKHQNHR-----AAPF-GIAAEYAHNEVALLGLG 201
QY 189 -QAVPHYALFIVPIHTHTHGLFMEATW-----TANHD-----CIHGMWMPV-M 233
DB 202 TVGIPYIWCITGNLH-----LFTVSIWIIILFLQAVDHSGEYFPWSIH-NFLPWA 253
QY 234 GAGYHTIHTTNYKHNHYTYIMDMWFGSLRDLLE---EDDNKDSFKAE 281
DB 254 GADHDEHHNYETIGGYSSFRWMDFLDTEAGPRAKKGREKXVQNVEKLD 304

RESULT 9
SUR2_YEAST STANDARD; PRT; 349 AA.
ID SUR2_YEAST
AC P38992;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE SUR2 PROTEIN (SYRINGOMYCIN RESPONSE PROTEIN 2).
GN SUR2 OR SYR2 OR YDR297W OR D9740.8.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
ON NCBI_TaxID=4932;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K21-1C;
RA Desjardes L., Durens P., Juguelin H., Cassagne C., Bonneau M.,
RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K21-1C;
RA Clifton P., Wang Y., Mochizuki D., Miyakawa T., Wangspa R.,
RT Hughes J., Takemoto J.Y.;
RT "SYR2, a gene necessary for syringomycin growth inhibition of
RT Saccharomyces cerevisiae."
RL Microbiology 142:477-484(1996).
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favelli A., Fulton L., Gattung S., Greco T., Kirsten J.,
RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA Madis E., Meneses S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Riken L., Riles L., Taich A., Trevasakis E., Vignati D.,
RA Wilcox L., Wohldman P., Vaudin M., Wilson R., Waterston R.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN THE RESPONSE TO SYRINGOMYCIN.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U07171; AAA16608.1; -
DR EMBL: U10427; AAB41115.1; -
DR EMBL: U28374; AAB64733.1; -
DR PIR: S48533; S48533.
DR SGD: S0002705; SUR2.
DR InterPro: IPR001541; Sterol_desat.
DR Pfam: PF01598; Sterol_desat; 1.
KW Transmembrane.
FT TRANSMEM 9 29 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 148 168 POTENTIAL.
FT TRANSMEM 209 229 POTENTIAL.
SQ SEQUENCE 349 AA; 40734 MW; 406D778092CA75C1 CRC64;

Query Match 7.58; Score 117; DB 1; Length 349;
Best Local Similarity 21.9%; Pred. No. 0.028;
Matches 66; Conservative 55; Mismatches 114; Indels 66; Gaps 17;

QY 3 ADNVLQMGVDETSP-----YNRIVLSHLRPNIMRPLRHLQDTWLRNLAGTLLYF- 57
DB 9 AAGSPFLAFGLKTSFGENHYAKAPAINLRK---ESLRLPMSDGLVALVAVVAWALSG 65
QY 58 FLWCFYIYVL--KINVYLPRDAIPTIKAMRLQMEVAMKAMPWYTLPTVSESMT----- 109
DB 66 IFHYIDTFLHAEKRYRIHSEEVAKRNKASRMHVEFY-----ILQHIQIIVGLIMH 118
QY 110 -----ERGWTKCFASIDE-----FGWILYVYTAIY--LVVEFGIYMHRE 149
DB 119 FEPIYMGFEENAMWKKLRADLRPIITDAIYGYWYGSAALKIFAGFLVDPTWOYELHRL 178
QY 150 LHDIKPLKYYIATHATHTIYNKONTLSPFA-GLAF-HPVCG-ILQAVPHYALFIVPIHTT 206
DB 179 MHMNTLTLMKTHSVNH-----ELIYPIYAGALFNPNVGVFLDITIGTIAHTLTHIRE 233
QY 207 HIGLFEATWNTANTHDCIHGNIMPV-----MGAGYHTIHTTY--KHNVGH--YTI 254
DB 234 QILLETFAIMKTVDH-C--GYALPLDPQWLFPRNNAVYHHIHQFOIKTNFQAPFTFE 290
QY 255 W 255
DB 291 W 291

RESULT 10
YCHO_YEAST STANDARD; PRT; 615 AA.
ID YCHO_YEAST
AC P25596; P25597; P25599; P87002;
DT 01-MAY-1992 (Rel. 22, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)

```
DE HYPOTHEICAL 68.9 KDA PROTEIN IN HML 5' REGION.  
NC YCLO073C OR YCL070C/YCL071C/YCL073C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OX Saccharomycetales; Saccharomycetaceae; Saccharomycos.  
RN NCBL_taxid=4932;  
RP  
RA SEQUENCE FROM N.A., Steensma H.V.;  
van der Aart O.T.M., Steensma H.V.;  
RL Submitted (MAR-1992) to the EMBL/GenBank/DDBJ databases.  
RR [2]  
RS REVISIONS.  
RW Gromadka R.;  
RA Submitted (JAN-1996) to the EMBL/GenBank/DDBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: BELONGS TO THE YCL070C/YHL047C/YKR106W FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL; X59720; CAA42397.1; -.  
SD SGD; S0000575; YCL073C.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 59 79 POTENTIAL.  
FT TRANSMEM 121 141 POTENTIAL.  
FT TRANSMEM 153 173 POTENTIAL.  
FT TRANSMEM 187 207 POTENTIAL.  
FT TRANSMEM 217 237 POTENTIAL.  
FT TRANSMEM 276 296 POTENTIAL.  
FT TRANSMEM 308 328 POTENTIAL.  
FT TRANSMEM 344 364 POTENTIAL.  
FT TRANSMEM 384 404 POTENTIAL.  
FT TRANSMEM 409 429 POTENTIAL.  
FT TRANSMEM 441 461 POTENTIAL.  
FT TRANSMEM 472 492 POTENTIAL.  
FT TRANSMEM 548 568 POTENTIAL.  
SQ SEQUENCE 615 AA; 68868 MW; D628405FFDC9F12 CRC64;
```

DT	01-JUN-1994 (Rel. 29, Created)
DT	01-JUN-1994 (Rel. 29, Last sequence update)
DT	01-FEB-1996 (Rel. 33, Last annotation update)
DE	HYPOTHEtical 69.0 KDA PROTEIN IN SIR1 3' REGION.
DE	YKR106W.
OS	Saccharomyces cerevisiae (Baker's yeast).
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycos.
OX	NCBI_TaxID=4932;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Gallion L., Dujon B.;
RL	Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC	-1- SIMILARITY: BELONGS TO THE YCL70C/YHL047C/YKR106W FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on ways
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/)
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; Z28202; CAA82047.1; -.
DR	PIR; S40624; S40624.
DR	SCD; S0001814; YKR106W.
KW	Hypothetical protein; Transmembrane.
FT	TRANSMEM 59 79 POTENTIAL.
FT	TRANSMEM 121 141 POTENTIAL.
FT	TRANSMEM 153 173 POTENTIAL.
FT	TRANSMEM 217 237 POTENTIAL.
FT	TRANSMEM 276 296 POTENTIAL.
FT	TRANSMEM 308 328 POTENTIAL.
FT	TRANSMEM 344 364 POTENTIAL.
FT	TRANSMEM 384 404 POTENTIAL.
FT	TRANSMEM 408 428 POTENTIAL.
FT	TRANSMEM 441 461 POTENTIAL.
FT	TRANSMEM 472 492 POTENTIAL.
FT	TRANSMEM 549 569 POTENTIAL.
SEQU	SEQUENCE 615 AA; 68954 MW; B548FBC6E9326AD CRC64;


```

Db 224 AACLSSEFGGLIAGCSKISGLKDMQYI-YIVEGICISLGFVPYAFGLSKNLEDSWEPN 282
QY 154 KPLKYIAHATHHHYKNTLSP-----FAGLAHPVD---GLQA 190
Db 283 KEKEYISER--YKTNATDPEPEKEMFQVQAVKDKTWASAVALEGIDLTTEGLTFV 339
QY 191 VPHVIA-----LFIPIHPTTHIGLLEMEAITANIH-----DCIGHNIM-- 230
Db 340 LPIITSMGFTNRAQMLTPYIFRLAI-VFICAVMSDRKRLSPFLGACLTTSIGIA 398
QY 231 PVNAGAYHTIHTTYKHNHYGTIAMD-----WMFGS 262
Db 399 IVLGSQVHGVR-----FGVYILCMGIYVNAACNMLLSGN 434

RESULT 14
YRR6_MYCA STANDARD; PRT; 279 AA.
ID YRR6_MYCA
AC P43045;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE HYPONHEICAL 33.0 KDA PROTEIN IN LICA 3' REGION (ORF R6).
OS Mycoplasma capricolum.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Mollicutes;
  Entomoplasmatidae.
OX NCBI_TaxID=2095;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 27343;
RX MEDLINE=94051609; Pubmed=8233831;
RA Miyata M., Sano K.-I., Okada R., Fukumura T.;
RT "Mapping of replication initiation site in Mycoplasma capricolum
  genome by two-dimensional gel-electrophoretic analysis.";
RL Nucleic Acids Res. 21:4816-4823(1993).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D14982; BAA03623.1;
KW Hypothetical protein.
SQ SEQUENCE 279 AA; 33031 MW; D6D1A265B88D6C18 CRC64;

Query Match 5.8%; Score 90.5; DB 1; Length 279;
Best Local Similarity 22.7%; Pred. No. 2.4; Mismatches 76; Indels 85; Gaps 14;
Matches 58; Conservative 37;

QY 52 LYFIS--GFLW-CFYIYLKINVLPR-----DAIPTIKAMRLQMFVAKA 94
Db 26 LFLSLIGFIFTSYTTDKLAIFEFKGEYIRVMSIFDIIGTTELMFLEFNIMVL 85
QY 95 M-PWYTLPLPVSESMIRGWTGCFASIDFEGMILFYVIAIYLVFVEGGIYWMHRELHDI 153
Db 86 IESWFL--KSKTKDNFMFK-----NMKILKVIYIVYVVEVV-----I 123
QY 154 KPLKYI-----NATHHIYNKONTLSPFAGLAHPV-----DG 186
Db 124 KCIITTYKINADNGFGGADIVLLSKYRNICLIVSLVHICIGLENGEYIIHKFNKNDP 183
QY 187 I-----LQAVPVIALFIYPIHPTTHIGLLEMEAITANIHDCIHNIM--PVNAGAY- 237
Db 184 IYLDKXWIAQV--ILFIYILSYTIIYVLKGMTS--RPYYNIYIYDLDLKQVKNLGNH 238
QY 238 ----HTIHTTYKHNHY 249
Db 239 DMVDHYLNOSTFKHGF 254

```

```

RESULT 15
COX3_MARPO STANDARD; PRT; 265 AA.
ID COX3_MARPO
AC P26858;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-1999 (Rel. 38, Last annotation update)
DE CYTOCHROME C OXIDASE POLYPEPTIDE III (EC 1.9.3.1).
GN COX3 OR COXIII.
OS Marchantia polymorpha (Liverwort).
OG Marchantia.
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Marchantiophyta;
  Marchantiales: Marchantiaceae: Marchantia.
OX NCBI_TaxID=3197;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92114051; Pubmed=1731062;
RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,
  Akashi K., Kanegae T., Ogura Y., Kohchi T., Ohyama K.;
RT "Gene organization deduced from the complete sequence of liverwort
  Marchantia polymorpha mitochondrial DNA. A primitive form of plant
  mitochondrial genome.";
RL J. Mol. Biol. 223:1-7(1992).
CC -1- FUNCTION: SUBUNIT I, II, AND III FORM THE FUNCTIONAL CORE OF
  -1- THE ENZYME COMPLEX.
CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
  4 FERRICYTOCHROME C.
CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 3 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M68929; AAC09434.1;
DR PIR: S25951; S25951.
DR HSSP: P00415; 10CC.
DR Mendel: 2057; MARPO:cox3.1.
DR InterPro: IPR000298; Cytochrome_c.
DR Pfam: PF00510; COX3; 1.
DR PROSITE: PS50253; COX3; 1.
KW Oxidoreductase; Mitochondrion; Transmembrane.
SQ SEQUENCE 265 AA; 29595 MW; 43DIC7820C7E7BC CRC64;

Query Match 5.8%; Score 90; DB 1; Length 265;
Best Local Similarity 23.1%; Pred. No. 2.5; Mismatches 53; Indels 62; Gaps 9;
Matches 42; Conservative 25;

QY 123 FGLIYEV-----YIATLVF-----VEFGIYWMHRELHDIKPL-YKYLHA----- 162
Db 84 YGIIIFIVSEVMFLAFWAFHSSLAFTYELGAIWPKGIVSDPMGIPLNLTILLSS 143
QY 163 -----THHIYNKONTLSPFAGLAHPVDGILLQAVPVIALFIYPIHPTTHIGLLEMEAI 216
Db 144 GAAVTWANH-----AILAGLR-----QNAVYALIAVFLALVETGFGGLEIYIAP 188
QY 217 WTANIHDCIHNIMPVNAGAYHTI-----HTTYKHNHYGTIAMD 258
Db 189 FT--ISDGIYSTF-FLATGFHGFHVIIGTIFLLICGIRQYLGHTTPKIHGFENAAFYW 245
QY 259 MF 260
Db 246 HF 247

Search completed: April 4, 2002, 17:08:08

```

Mon Apr 8 17:54:19 2002

us-09-775-879-21.rsp

Page 10

Job time: 101 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 17:05:57 ; Search time 16.4 Seconds
(without alignments)
1305.186 Million cell updates/sec

Title: US-09-775-879-21
Perfect score: 1562
Sequence: 1 MADNAYLMQFVDETSFYNR.....SLRDPLEDDNKRSEFKAE 281

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_68:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1538	98.5	281	2 S71251	C-5 sterol desatur
2	304.5	19.5	329	2 T40027	hypothetical C-5 S
3	295.5	18.5	365	2 J01146	C-5 sterol desatur
4	279.5	17.9	300	2 T37759	C-5 sterol desatur
5	174.5	11.2	286	2 T22444	hypothetical prote
6	167	10.7	253	2 G84695	probable C-4 stero
7	152.5	9.8	231	2 S76372	hypothetical prote
8	152.5	9.8	309	2 S64354	ERG25 protein - ye
9	149.5	9.6	258	2 B96718	probable sterol de
10	149	9.5	300	2 T16255	hypothetical prote
11	147.5	9.4	269	2 T22443	hypothetical prote
12	146	9.3	300	2 T38986	probable C-4 methy
13	138	8.8	303	2 T06618	hypothetical prote
14	129.5	8.3	258	2 A86277	hypothetical prote
15	127	8.1	820	2 T04570	hypothetical prote
16	122	7.8	255	2 F83409	hypothetical prote
17	117	7.5	293	2 T40740	hypothetical integ
18	117	7.5	349	2 S48533	SUR2 protein - yea
19	115.5	7.4	635	2 T02536	CER1-like protein
20	115	7.4	246	2 S74961	hypothetical prote
21	113	7.2	304	2 D82189	conserved hypothet
22	102.5	6.6	261	2 T12402	cytochrome-c oxida
23	99	6.3	493	2 T18789	hypothetical prote
24	98.5	6.3	301	2 G83556	hypothetical prote
25	98	6.3	615	2 S74277	probable membrane
26	98	6.3	615	2 S40624	probable membrane
27	94	6.0	555	2 T04166	glossyl homolog -
28	93.5	6.0	386	1 S41691	ubiquinol--cytochr
29	93	6.0	936	2 B64567	cytochrome c bioge

30	92.5	5.9	542	2 B81662	apolipoprotein N-a
31	92.5	5.9	959	2 T25704	hypothetical prote
32	91	5.8	534	2 S64593	probable membrane
33	90.5	5.8	279	2 S42125	hypothetical prote
34	90	5.8	265	2 S25951	cytochrome-c oxida
35	90	5.8	318	1 S75765	stearoyl-CoA desat
36	89	5.7	351	2 T21004	hypothetical prote
37	89	5.7	373	2 G85355	nodulin-like prote
38	89	5.7	649	2 B38129	bo-type ubiquinol
39	88.5	5.7	308	2 S22928	ubiquinol--cytochr
40	88.5	5.7	353	2 T33974	hypothetical prote
41	88.5	5.7	1188	2 S48861	gene e1 protein -
42	88	5.6	321	2 A85970	probable transport
43	88	5.6	694	2 T12712	NADH dehydrogenase
44	88	5.6	694	2 T12675	NADH dehydrogenase
45	87.5	5.6	261	2 T11525	cytochrome-c oxida

ALIGNMENTS

RESULT 1
S71251
C-5 sterol desaturase (EC 1.-.-.-) - Arabidopsis thaliana
N: Alternate names: delta7-sterol-C5-desaturase; sterol-C5-desaturase
C: Species: Arabidopsis thaliana (mouse-ear cress)
C: Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 17-Mar-2000
C: Accession: S71251
R: Gachotte, D.; Husselstein, T.; Bard, M.; Benveniste, P.
submitted to the EMBL Data Library, July 1995
A: Description: Cloning a plant sterol-C5-desaturase by functional complementation of
A: Reference number: S71251
A: Accession: S71251
A: Molecule type: mRNA
A: Residues: 1-281 <GAC>
A: Cross-References: EMBL: X90454; NID: g1061037; PID: g106138
A: Superfamily: Saccharomyces cerevisiae ERG25 protein
C: Keywords: oxidoreductase

Query Match 98.5%; Score 1538; DB 2; Length 281;
Best Local Similarity 98.6%; Pred. No. 2,7e+128;
Matches 277; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MADNAYLMQFVDETSFYNRIVSHLLPANLWEPRLPHLOTWLRNYAGTLTYISGFLM 60
DB 1 MADNAYLMQFVDETSFYNRIVSHLLPANLWEPRLPHLOTWLRNYAGTLTYISGFLM 60

QY 61 CFYIYKINVLPRKDAIPTIKAMRLQMFVAMKAMPYTLPTVSE;MIERGWTGCFASI 120
DB 61 CFYIYKINVLPRKDAIPTIKAMRLQMFVAMKAMPYTLPTVSE;MIERGWTGCFASI 120

QY 121 DEFGWILYFVYLAIVYVEFGIYWMHRELHDIKPLKYLLHATNHT(NKONTLSPPAGLA 180
DB 121 GEGWILYFVYLAIVYVEFGIYWMHRELHDIKPLKYLLHATNHT(NKONTLSPPAGLA 180

QY 181 FHHVQSILOAVPHVIALFTVPIHFTTHIGLFPHEALWTANIHCHISNIMPVAGAGYHTI 240
DB 181 FHHVDSIILOAVPHVIALFTVPIHFTTHIGLFPHEALWTANIHCHISNIMPVAGAGYHTI 240

QY 241 HHTYKHNKGHTYIMDMFGSLRDLPLEDDNKRSEFKAE 281
DB 241 HHTYKHNKGHTYIMDMFGSLRDLPLEDDNKRSEFKAE 281

RESULT 2
T40027
hypothetical C-5 sterol desaturase - fission yeast (Schizosaccharomyces pombe)
C: Species: Schizosaccharomyces pombe
C: Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C: Accession: T40027
R: Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lauber, J.; Hilbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, December 1997

Db 159 CTPTASHAFKSDAGFLOSILPYHLFPFPLKLTALFTFVNWFSIMIHOCKYISNNPI 218
 OY 233 M-GAGYHTIIHHTTYKHNHYGHTIMDMMGSLRDP 266
 Db 219 INGAHHNGHHIYFVNYGQFTTFLDRLGNSFRAP 253

RESULT 5

T22444

hypothetical protein F49E12.10 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000

C:Accession: T22444

R:Thomas, K.

submitted to the EMBL Data Library, October 1995

A:Reference number: Z19565

A:Accession: T22444

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-286 <FILL>

A:Cross-references: EMBL:Z66520; PIDN:CA91384.1; GSPDB:GN00020; CESP:F49E12.10

A:Experimental source: clone F49E12

C:Genetics:

A:Gene: CESP:F49E12.10

A:Map position: 2

A:Introns: 21/2; 130/2; 153/1; 221/2

C:Superfamily: Saccharomyces cerevisiae ERG25 protein

Query Match 11.2%; Score 174.5; DB 2; Length 286;
 Best Local Similarity 22.4%; Pred. No. 4.3e-08;

Matches 64; Conservative 49; Mismatches 112; Indels 61; Gaps 14;

OY 39 LQTLRLNLAFT--LLYFISG-----FLKCFYIYLLKINYLPRDAIPTKAMRLQMFV 90

Db 5 LYRLLYNFDEDEMYLLVYVNAVAGISFWLNFILIIDVDPK---WQPYKIOEEK 60

OY 91 AMKAMPWYTLPTVSESM-----IERGWTCKFASIDFGMTLVYVYIAIYLV 137

Db 61 KPSISKYISILKVGPNOLITPTVITLTFYVARMWGMDFGVIP-SWITLRLDACLCA 119

OY 138 FVEFGIWMARELDIKPLKYULNATHNIYKONTLSFPAAGLAFHPVDGILQAVPHVIAL 197

Db 120 MDEIGFYTHRLFHNPK-LYKHKKHHEWNAVVIS---SIYAPLE--HAISNLSPI 172

OY 198 FIVYIHFTTHI---GLFMEAIWTANIHDCIHGNITPVM-GAGYHTIHTTYKHNHYGT 253

Db 173 YLGAVLFCHVSHYIFTSVAIILTTFFHH--SGYHFPMLSAEHDFHKKVFNCCYGGP 230

OY 254 IWMQWNG-----SLRDP-----ILLEEDONKDSFKK 279

Db 231 --LWMLHGTDTFKRSIHPNDIYVYGTTPIKELIPEDKNNNKK 274

RESULT 6
 G84695
 probable C-4 sterol methyl oxidase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: G84695

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffett, K.S.; Crokin, L.A.; Shen, M.; Vankar, S.E.; Umayam, L.; Tallon, L.;

euser, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402:761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487

A:Accession: G84695

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-253 <STO>

A:Cross-references: GB:AE002093; MID:g3980396; PIDN:AC95199.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g29390
 A:Map position: 2

Query Match 10.7%; Score 167; DB 2; Length 253;

Best Local Similarity 22.2%; Pred. No. 1.7e-07;

Matches 58; Conservative 32; Mismatches 75; Indels 96; Gaps 11;

OY 52 LYFISGLMCF-----YIYLLKINYLPRDAIPTK 82

Db 26 VFFLSQLPYIFLERTGELSNYKIQTKSNTPPEAGKCIARLLHCCVNLPLMASYPVTR 85

OY 83 ARRLQFVAMKAMPWYTLPTVSESMIERGWTCKFASIDFGMTLVYVYIAIYLVVERG 142

Db 86 FMGME-----SSFP-----LPS-----WKVVSQILFFIIEDEV 115

OY 143 IYWMARELDIKPLKYULNATHNIYKONTLSFPA-----GLAFHPVDGILQAV 190

Db 116 FYWGRIRLH-TKWLKYNVSHVHEX-----APFGLTSEVAPAEILVIGFATVGPALT 169

OY 191 VPHVIALFVPIHFTTHIGLLEMEAIWTANIHDCIH-----GNIPVNI-GAGYHTIHTH- 243

Db 170 GPHLLITLW-----WMLRVLETVEAHGCHYHFPWSPSNFLPL:GGADPHDYHNR 219

OY 244 --TYKHNHYGTIMDMMGFS 262

Db 220 LYTKSGNYSSTFYVMDMIRGT 240

RESULT 7

S76372

hypothetical protein - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999

C:Accession: S76372

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,

O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys

s.

A:Reference number: S74322; MUID:97061201

A:Accession: S76372

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-231 <KAN>

A:Cross-references: EMBL:D64000; GB:AB001339; MID:g1001464; PIDN:BA10224.1; PID:d101

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 9.8%; Score 152.5; DB 2; Length 231;

Best Local Similarity 23.6%; Pred. No. 3e-06;

Matches 58; Conservative 42; Mismatches 113; Indels 33; Gaps 9;

OY 23 LSHLPANLMEPLPHLQTLNRYL---AGTLTYFISGLMCFYIYLLKINYLPRDAIP 79

Db 1 MDHL-----FWIYLVFFAILLFRFLIAGTYAFTSPGOSILKNHLPLN-----IP 49

OY 80 TKAMRLQFVAMKAMPWYTLPTVSESMIERGWTCKFASIDFG--NIIYVYIAIYLVF 138

Db 50 SSOSIRKDKILSIISAIIPLAGAFITSSYGKMNLVDPYOHGNYIGISYVL-LFL 108

OY 139 VEFGIWMARELDIKPLKYULNATHNIYKONTLSFPAAGLAFHPVDGILQAVPHVIALF 198

Db 109 QDTYVFYTHRLFHNPK-LFSLFKHGHLSRYPLTFSF--AFDLPZALVQSFLLIYVS 164

OY 199 IYPIHFTTHIGLLEMEAIWTANIHDCIH-----GNIPVNI-GAGYHTIHTTYKH 247

Db 165 LIRLHFTTILALIMKTIMSVINHLCIDRLPLAFPHIMLGKTF--ITVANHSLHLKYNA 222

OY 248 NYGHTY 253

Db 223 NYGLYS 228

RESULT 8
S64354
ERG25 protein - yeast (*Saccharomyces cerevisiae*)
N/Alternate names: protein G4358; protein YGR060w
C/Species: *Saccharomyces cerevisiae*
C/Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 20-Jun-2000
C/Accession: S64354
R:Entian, K.D.; Rose, M.; Koetter, P.; Roehmer, A.; Sehmsam, I.; Hempel, S.
submitted to the Protein Sequence Database, May 1996
A/Reference number: S64335
A/Accession: S64354
A/Molecule type: DNA
A/Residues: 1-309 <ENT>
A:Cross-references: EMBL:Z72845; NID:q1323076; PIDN:CA97062.1; PID:q1323077; MIPS:YGR060w
A:Experimental source: strain S288C
C/Genetics:
A:Gene: SGD:ERG25
A:Cross-references: SGD:S0003292; MIPS:YGR060w
A:Map position: 7R
C:Superfamily: *Saccharomyces cerevisiae* ERG25 protein
C:Keywords: Transmembrane protein
:51-67/Domain: Transmembrane #status predicted <TM>

Db 179 LHNHEFELEFVOTFTTTLIP-----WIFPHCLCTIWIWFFLIQSYSEYHIGIDYDEPPALHRI 233

Qy 229 IMPVMGAGYHTTIHH 242

Db 234 FWFYSGAPAHDMHH 247

RESULT 11
T22443
hypothetical protein F49E12.9 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #extl_change 02-Jun-2000
C:Accession: T22443
R:Thomas, K.
submitted to the EMBL Data Library, October 1995
A:Reference number: Z19565
A:Accession: T22443
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-269 <WIL>
A:Cross-references: EMBL:Z66520; PIDN:CAA91383.1; GSPDB:GN00020; CESP:F49E12.9
A:Experimental source: clone F49E12
C:Genetics:
A:Gene: CESP:F49E12.9
A:Map position: 2
A:Introns: 20/2, 96/1, 129/2, 152/1, 174/1, 202/1, 220/2
;Superfamily: Saccharomyces cerevisiae ERK25 protein

Query Match	9.4%	Score 147.5	DB 2	Length 269
Best Local Similarity	23.2%	Pred. No. 9	7=06	
Matches	55	Conservative	40	Mismatches 99
			Indels 43	Gaps 11

QY	51	LVYISGFL-----WCFYIYLKINVLPRDAIP-----TIKARLDM	88
		: : : : : : : : :	
Db	18	LIYVIGGIVAININFWLNFLEFFITIDMDIPWVCQYKIDQEKRPRLSYGAFKAVILANQ	77
QY	89	FVAKKAMPWYTLPLPVSESMIERGM-TKCPASIDFEGMIVYVYLAIVLPAEFGIYMMH	147
		: : : : : : : : :	
Db	78	FIAGPLITLFWFPAV-----WFGSFTGRLPSGMOI-FRDIIVISLICEIGIFYSH	128
QY	148	RELHDIKPLVYKYLATNHIYNKONTLSBPAGIAFHPRVGDILOAVENIALFTVPIHFTTH	207
		: : : : : : : : :	
Db	129	RLFNHPK-IKYKIKKHNHEMTAPVST--SYCPRLP--HAISNLSVLLGPTICGSH	181
QY	208	IGLGEMLAIWANAINDCIH-GNIWYVM-GAGNTHIHHTYKKNYGHYITIMDMFMGCS	262
		: : : : : : : : :	
Db	182	VITLWIKMSIALISTGSHSGHNEFFMLSPERHDNHNKVVFNCFG--TGLDMHIGT	236

RESULT 12
T38986
probable c-4 methyl sterol oxidase - fission yeast (*Schizosaccharomyces pombe*)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
C:Accession: T38986
R:McGoually, R.C.; Rajandream, M.A.; Barrell, B.G.; Zimmermann, W.; Wambutt, R.
submitted to the EMBL Data Library, August 1999
A:Reference number: Z21822
A:Accession: T38986
A:Status: preliminary; translated from GB/EMBL/DDBT
A:Molecule type: DNA
A:Residues: 1-300 <EMC>
A:Cross-references: EMBL:AL109832; PIDN:CAB52730.1; GSPDB:GN00066; SPDB:SPAC630.08c
A:Experimental source: strain 972h; cosmid c630
C:Genetics:
A:Gene: SPDB:SPAC630.08c
A:Map position: 1
A:Superfamily: Saccharomyces cerevisiae ERG25 protein

Query Match	9.38;	Score 146;	DB 2;	Length 300;
Best Local Similarity	23.68;	Pred. No. 1.5e-05;		

Matches 73; Conservative 41; Mismatches 113; Indels 82; Gaps 19;

```

QY      8 LMOVDSTSYNNRYLVSHTLRYNMLBRLPHFLOQTYLRNLTAGTLVYFI3GFLWMCYIYLL 67
Db      20 LAQNHRLPRLNVEQDMLAY-----YKMDNDVYATGL--MSFLHBLITFG 62

QY      68 KINUYLRKDAIPTIKARLO-MFVANKAMPV-YTLRLTVSESME--RSM--TKC----- 116
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      63 RCIMWMLIIDAMPFRMRKIQPKKVPRLAQWECTRLVLSHFTVELPQYMLDRPCATFG 122

QY      117 -----FASIDFGWILYENYAIYLVVEWEGUYMNHRELNDIKRLYKYLHATNHIYKQ 170
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      123 LSTSVPRPPVTKMIV-----QITLFFELBEDTWMYAHRLFH-YGIFYEFIFKHVHARYS-- 174

QY      171 NTLSPFGFLA---FHPVDGILQA-----VPHYALFIVYIHFTT---IGLLFMAFI--- 216
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      175 ---APF-GLSAEYAHPLLEILLAGTVPVRLMWCYETHDLHLYTMXIVITELTFQAVDSH 230

QY      217 -----WTANILHDCIHGNIMPV-MGAGYUPTIHNTVYKKNHYGYNTIYMLIMFQSLDPLLE 269
        : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      231 AGYDFPWSLN-----KFLPIWAGADHNDHYHMAKDNFSSSFRV--W-----DAVLK 275

QY      270 EDDNKDSDKF 278
        : : : : :
Db      276 TDQNTYHQFK 284

```

RESULT 13
T06618
hypothetical protein F16J13.180 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 10-Dec-1999
C:Accession: T06618
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft
submitted to the Protein Sequence Database, April 1999
A:Reference number: Z15789
A:Accession: T06618
A:Molecule type: DNA
A:Residues: 1-303 <BEV>
A:Cross-references: EMBL:AL049638; GSPDB:GN00062; ATSP:F1J13.180
A:Experimental source: cultivar Columbia; BAC clone F16J13
C:Genetics:
A:Gene: ATSP:F16J13.180
A:Map position: 4
A:Insertion: 112/3; 220/2; 269/3; 285/3
A:Superfamily: Saccharomyces cerevisiae ERG25 protein

Query Match	8.8%;	Score 138;	DB 2;	Length 303;
Best Local Similarity	23.5%;	Pred. No. 7.6e-05;		
Matches	52;	Conservative	33;	Mismatches 72;
				Infels 64;
				Gaps 11;

```

QY 86 LOMFLVAKAMMYATLLPVSSESMTERGCTKCFASIDEGMILYRVUATYAIYUVEVEGFIY 145
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 111 IOMELINSGLP-----LPITBMSO-----LWFFLELDYNNW 145
QY 146 MHRELNDIKPLUKYUUAHTNHNINNO--NLSPEFACLAEPVNDGILQAVPHVATLEIVPIHF 204
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 146 VHRFHHS-KWUYDKIHNHNHETARIGAAUYA----HMVEVLLLGIRTEGRAIAGHM 200
QY 205 TTH---IGLLEMEAI-----WTANIDHCIGNITPVWAGAYHTIHNT--YKNY 249
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 201 ITFWMILALROMEMAIETHSGYDPMSPFKYIRPYG-----GAYUHEYNHVUGOSOSNF 254
QY 250 GHYITIMDMWMS-----LRDPLLEDDDKKOSFKKA 280
      :| | | | | :| | | | | :| | | | | :| | | | | :| | | | | :| | | | |
Db 255 ASVFUYCDYIGTDKMWLYOGYRQKRLLEO--IKESKSS 293

```

RESULT 14
A86277
hypothetical protein AAF43928.1 [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 17:05:32 ; Search time 12.63 Seconds
(without alignments)
500.667 Million cell updates/sec

Title: US-09-775-879-21

Perfect score: 1562

Sequence: 1 MAADNAYLMQFVDETSFYNR.....SLRDLLEEDDNKDSFKAE 281

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTUTS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	135.5	8.7	625	US-08-581-148C-18	Sequence 18, Appl
2	104.5	6.7	594	US-08-581-148C-16	Sequence 14, Appl
3	98	6.3	555	US-08-581-148C-16	Sequence 16, Appl
4	90	5.8	318	US-08-309-182B-3	Sequence 3, Appl
5	85	5.4	407	US-08-742-440A-3	Sequence 3, Appl
6	81	5.2	1732	US-08-477-451-14	Sequence 14, Appl
7	81	5.2	2713	PCR-US96-01735-1	Sequence 1, Appl
8	80.5	5.2	408	US-08-742-440A-6	Sequence 6, Appl
9	80	5.1	360	US-08-597-236-11	Sequence 11, Appl
10	80	5.1	360	US-08-746-682A-11	Sequence 11, Appl
11	76.5	4.9	325	US-08-467-947A-29	Sequence 29, Appl
12	76.5	4.9	325	US-08-467-947A-29	Sequence 29, Appl
13	76.5	4.9	325	US-08-988-876-8	Sequence 8, Appl
14	76.5	4.9	338	US-09-303-524A-2	Sequence 2, Appl
15	76.5	4.9	439	US-08-716-873-4	Sequence 4, Appl
16	76.5	4.9	439	US-09-368-431-4	Sequence 4, Appl
17	75.5	4.8	355	US-08-461-244-2	Sequence 2, Appl
18	75.5	4.8	355	US-09-045-583-56	Sequence 56, Appl
19	75.5	4.8	362	US-08-902-853-5	Sequence 5, Appl
20	75.5	4.8	380	US-08-244-205-11	Sequence 11, Appl
21	75.5	4.8	380	US-08-244-205-11	Sequence 11, Appl
22	75.5	4.8	386	US-08-244-205-2	Sequence 2, Appl
23	75.5	4.8	386	US-08-244-205-2	Sequence 2, Appl
24	75.5	4.8	1581	US-08-726-320-3	Sequence 3, Appl
25	75.5	4.8	1581	US-09-208-716-3	Sequence 3, Appl
26	75	4.8	438	US-08-952-365-2	Sequence 2, Appl
27	75	4.8	3165	US-08-459-146-3	Sequence 3, Appl

28	75	4.8	3165	US-08-459-065-3	Sequence 3, Appl
29	75	4.8	3287	US-08-477-451-7	Sequence 7, Appl
30	74	4.7	261	US-09-097-889-19	Sequence 19, Appl
31	73.5	4.7	439	US-08-716-873-2	Sequence 2, Appl
32	73.5	4.7	439	US-09-368-431-2	Sequence 2, Appl
33	73.5	4.7	479	US-08-484-105-10	Sequence 10, Appl
34	73.5	4.7	479	US-08-484-105-10	Sequence 10, Appl
35	73	4.7	474	US-08-453-742-2	Sequence 2, Appl
36	73	4.7	474	US-08-454-464-2	Sequence 2, Appl
37	73	4.7	474	US-08-453-222-2	Sequence 2, Appl
38	73	4.7	474	US-08-452-802-2	Sequence 2, Appl
39	73	4.7	477	US-08-453-742-27	Sequence 27, Appl
40	73	4.7	477	US-08-454-464-27	Sequence 27, Appl
41	73	4.7	477	US-08-453-222-27	Sequence 27, Appl
42	73	4.7	477	US-08-452-802-27	Sequence 27, Appl
43	72.5	4.6	1498	US-08-404-531B-28	Sequence 28, Appl
44	72.5	4.6	1498	US-08-404-531B-29	Sequence 29, Appl
45	72.5	4.6	1498	US-08-476-900A-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-08-581-148C-18
; Sequence 18, Application US/08581148C
; Patent No. 6060644
; GENERAL INFORMATION:
; APPLICANT: Schnable, Patrick S.
; APPLICANT: Robertson, Donald S.
; APPLICANT: Hansen, Joel D.
; APPLICANT: Nikolau, Basil J.
; APPLICANT: Xu, Xiaojie
; APPLICANT: Xia, Yijl
; TITLE OF INVENTION: ISOLATION AND USE OF CUTICULAR LIPID
; TITLE OF INVENTION: GENES
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Volt & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/581,148C
; FILING DATE: 29-DEC-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Larcher, Carol
; REGISTRATION NUMBER: 35243
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 625 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-581-148C-18

Query Match 8.7%; Score 135.5; DB 3; Length 625;
Best Local Similarity 22.8%; Pred. No. 6.5e-07;
Matches 55; Conservative 34; Mismatches 71; Indels 81; Gaps 11;

QY 64 IYIKINIVYLPKDAIPITIKAMRLQMEVAMKAMPYTLTPVSESNIERGCTKCFASIDEF 123
DB 105 LEYIGIN-----LAEKGKOLPMWRDGVLMGALIHFG----- 136
QY 124 GWILYFYVIAIYLVFEGIYMMHRELHDIKPLKYKYLHATHH--IYNQNT--LSFAGL 179
DB 137 -----PVEFLYVWKALHH-HFLYSRYSHHSSIVTEPITSVHPFAE- 180
QY 180 AFHPVDGILQAVPHVIAL-----FIVPIHFTTHIGLFEAMIAWTANIHCICIG 227
DB 181 --HIAVFLFALPILTLTLVTKTASISFAGYIYIDFNMMNHCNFEILPKRLH----- 233
QY 228 NIMVVM-----GAGYHITHHTTKYKHNQYIIMDMMEGSL--RDLLEE-----DDNK 274
DB 234 -LFPPLKFLCYTPSYSHLHOFRTNYSLFMPLYDIYIGTDESTDTLVEKTERGDDRV 292
QY 275 D 275
DB 293 D 293

RESULT 2
US-08-581-148C-14
; Sequence 14, Application US/08581148C
; Patent No. 6060644
; GENERAL INFORMATION:
; APPLICANT: Schnable, Patrick S.
; APPLICANT: Robertson, Donald S.
; APPLICANT: Hansen, Joel D.
; APPLICANT: Nikolau, Basil J.
; APPLICANT: Xu, Xiaojie
; APPLICANT: Xia, Yijl
; TITLE OF INVENTION: ISOLATION AND USE OF CURRICULAR LIPID
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Leydig, Volt & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/581,148C
; FILING DATE: 29-DEC-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Larcher, Carol
; REGISTRATION NUMBER: 35243
; REFERENCE/DOCKET NUMBER: 71380
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 594 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-581-148C-14

Query Match 6.7%; Score 104.5; DB 3; Length 594;
Best Local Similarity 19.0%; Pred. No. 0.0021;
Matches 52; Conservative 34; Mismatches 74; Indels 113; Gaps 12;
QY 60 WCFPIY--YIKINIVYLPKDAIPITIKAMRLQMEVAMKAMPYTLTPVSESNIERGCTKCFASIDEF 111
DB 111

DB 9 WCFHIVISILRENLVM-----WYINICMMLFLTRNRRLHQ 45
QY 112 GWTKCFASID-EFGMIYFYV-----IAIYLVFVERG----- 142
DB 46 SID--FNQIDKEMWMDNVIYQALIASIATIMFQEFANLPVWKTKGLVAVIYHVYSE 103
QY 143 --IYMMHRELHDIKPLKYKYLHATHHYNKNTLSFAGLAFHPVDGILQAVPHVIALFTV 200
DB 104 PLYWMLRLHRT-----NYLTPYH-----SFHSSAVPQPVYVSTTFLE 144
QY 201 PIHFTTHIGL-----FMEAIWTANIHCICIG--NIMP----- 231
DB 145 ELLVAVGLPILCCSLSGYKSKSIYGYLVDFDLKGLCHSHNVEIMPHWIEFYVFFFRF 204
QY 232 -VMGAGYHTHTTKYKHNQYIIMDMMEGSL 263
DB 205 IITYPTIYSLHSEKSNVCLFMPLYDTMNTL 237

RESULT 3
US-08-581-148C-16
; Sequence 16, Application US/08581148C
; Patent No. 6060644
; GENERAL INFORMATION:
; APPLICANT: Schnable, Patrick S.
; APPLICANT: Robertson, Donald S.
; APPLICANT: Hansen, Joel D.
; APPLICANT: Nikolau, Basil J.
; APPLICANT: Xu, Xiaojie
; APPLICANT: Xia, Yijl
; TITLE OF INVENTION: ISOLATION AND USE OF CURRICULAR LIPID
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Leydig, Volt & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: US
; ZIP: 60601-6780
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/581,148C
; FILING DATE: 29-DEC-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Larcher, Carol
; REGISTRATION NUMBER: 35243
; REFERENCE/DOCKET NUMBER: 71380
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 555 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-581-148C-16

Query Match 6.3%; Score 98; DB 3; Length 555;
Best Local Similarity 21.5%; Pred. No. 0.01;
Matches 39; Conservative 27; Mismatches 59; Indels 56; Gaps 9;
QY 114 TKCFASIDEFGMIYFYVIAIYLVFEGIYMMHRELHDIKPLKYKYLHATHHYNKNTL 173
DB 50 TSDLSAMDRLGMA--IAVVLHVAVSEPAFYWAHRLH-LGPLRSRYSHLHNSFOATQAL 105

OY 174 SPFAGLAHPVDGI---LQAVPHVIALFVPIHFTTHIGLFMEAIWTANIDCHIGNI- 229
DB 106 T--AGFV--TPLESLLTLIVAMPHLOGL-----HGG-----TRLNELVYGHIS 144
OY 230 ---WFWGCA-----GHTHTHTTKYKINYGHTYIMDMWFGS 262
DB 145 SSTFVHGVOORGHJLQDFODFPEPLRYLYTPSYLSLHHRKSDNFCLEFMDLPALGCT 204
OY 263 L 263
DB 205 L 205

RESULT 4

US-08-309-182B-3
; Sequence 3, Application US/08309182B
; Patent No. 5639645
; GENERAL INFORMATION:
; APPLICANT: No. 563964510 MURATA
; TITLE OF INVENTION: A RECOMBINANT 9 DESATURASE AND A GENE
; TITLE OF INVENTION: ENCODING THE SAME
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 500 kb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/309,182B
; FILING DATE: September 20, 1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 318 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-309-182B-3

Query Match 5.8%; Score 90; DB 1; Length 318;
Best Local Similarity 19.0%; Pred. No. 0.039;
Matches 55; Conservative 38; Mismatches 117; Indels 80; Gaps 10;

OY 42 WIRNYLACTLVFISGFLMCFYIYLYKINVLPKDAIPTIK-----AM 84
DB 11 YLSKLFDSLSLVFNKRQLFRFVRFPMALPNDSPKPLTPAWTVIEFFTSIHVALLA 70
OY 85 RLOMVAMKAMPWYTLPLTVSESM--IERGWTGCA--SIDEQWILYFYIAIYL----- 136
DB 71 FLPOFFSKAAGMAFLLYVITGGITIGTFHRCISHRSFNVPKWLLEYFVICGLACOGG 130
OY 137 VVEVGIIWMHRELHDIPLYKYKILHATHIYNK----- 169

DB 131 VFEWVGLHRMHKFSDTTP-----DPHDSNKGFWMSHIGMMPEIPAKADIPRYKDI 183
OY 170 -----QNTL-----SPFAGLAHPVDGILOAVPHVIALFVPIHFTTHIGLFMEAIW 217
DB 184 QDDKRYQCQNNLLILOVALGLILFALGWPVINGIFRLVFEHPTVEVNSATHKFGY 243
OY 218 TANIDCHIGNIWPV---MGAGYHTIHTTKYKHNVGHTYI---MDYM 259
DB 244 VSHESNDYSRNCMWVALLTFGEGWNNH--AYQYARHOLQWVEVDLTMM 292

RESULT 5

US-08-742-440A-3
; Sequence 3, Application US/08742440A
; Patent No. 5892014
; GENERAL INFORMATION:
; APPLICANT: Coughlin, Shaun
; APPLICANT: Ishihara, Hiroaki
; APPLICANT: Connolly, Andrew
; TITLE OF INVENTION: Protease Activated Receptor
; TITLE OF INVENTION: 3 and Uses Thereof
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bozicevic & Reed, LLP
; STREET: 285 Hamilton Avenue, Suite 200
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,440A
; FILING DATE: 30-OCT-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: 36,677
; REFERENCE/DOCKET NUMBER: UCAL/060PAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-327-3400
; TELEFAX: 650 327-3231
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 407 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; US-08-742-440A-3

Query Match 5.4%; Score 85; DB 2; Length 407;
Best Local Similarity 23.7%; Pred. No. 0.2;
Matches 59; Conservative 33; Mismatches 85; Indels 72; Gaps 15;

OY 51 LLYFISGFLMCFYIYLYLKI-----NVLPRDAIPTIKAMRLOMFVAMKAMPWYTLAPT 103
DB 165 IAYHLNGNNWVGEWMCRTTVVFGNMYC---AIIILCMGGINRYLATAHPTYQKLPK 221
OY 104 VESMERGTWCSPASIDFGWILYFYIAIYLVFVSGIYMMHRL---HDIKPLKYKL 160
DB 222 RSFSLMCG-----IYMWVWFLYMLPEVILKQ--EYHLVHSEITTCHEVNVACESP 270
OY 161 HATHHIYNKQNTLSPFAGLAHPVDGILOAVPHVIALFVPIHFTTHI-GLLFMEAIWTA 219

DB 271 SSRFFY-----FVSLAF---GFL--IPEVITFC---YTLIHKKSKDRIMLG 313
QY 220 NIHDCI-----HGNIWPMGAGYHTIHTTQYKHNHYGHTIMDMWF-----GS 262
DB 314 YIRAVLLIVITFCFAPTNILV-----IHMANV---YHNHTDSLRYMILALCLGS 363
QY 263 LR--DPL 268
DB 364 LNSCLDPFL 372

RESULT 6

US-08-477-451-14
Sequence 14, Application US/08477451
Patent No. 5928865
GENERAL INFORMATION:
APPLICANT: Covacci, Antonello
TITLE OF INVENTION: Helicobacter Pylori CagI Region
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/477,451
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McClung, Barbara G.
REGISTRATION NUMBER: 33,113
REFERENCE/DOCKET NUMBER: 0335.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 510-601-2708
TELEFAX: 510-655-3542
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 1732 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-477-451-14

Query Match Best Local Similarity 5.2%; Score 81; DB 2; Length 1732;

Matches 43; Conservative 23; Mismatches 96; Indels 62; Gaps 7;

QY 75 KDAIPITKAMRLQMFVAMKAMPWYTLPTVSESIERGWTGCFASIDEFGWILYFYIAI 134
DB 1113 KAVIPTRKAIKDHCMIFLR-----RAMISNLMLDLLLFFLLIWIATFFH--- 1157
QY 135 YLVFVEGIGIWMHRELHDIDIPLYKYLH-----ATHIYKNTKNTLSPPAGLAFHPVDGI 187
DB 1158 ERCEFLVEDRMAVARISDCLMLVLLHFLLYOCOTRLHLVPGSYFLAPFLFFSSHKRKT 1217
QY 188 LQAVPHVIALFVIPIHTTHIGL-----LMEAI-----WTANIHDC 224
DB 1218 ER-----FYCCCEYEVNLHGVKTCYFLSLFLEFFLINGQEFFDPCLSNFTLTHNRC 1270
QY 225 IHGNIWPMGAGYH-----TIHTTQYKHNHYGHTIMDMW 258
DB 1271 LKTYTHFOFAFSTHQNAEQOVSKIHSPSSHSYHLYIMFOW 1314

RESULT 7

PCT-US96-01735-1

Sequence 1, Application PC/TUS9601735

GENERAL INFORMATION:

APPLICANT: Marks, Andrew R.

TITLE OF INVENTION: HUMAN T CELL INOSITOL 1,4,5,-TRISPHOSPHATE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond

STREET: 30 Rockefeller Plaza

CITY: New York

STATE: NY

COUNTRY: USA

ZIP: 10112-0228

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US96/01735

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/386,039

FILING DATE: 09-FEB-1995

ATTORNEY/AGENT INFORMATION:

NAME: Kote, Lisa B

REGISTRATION NUMBER: 35,225

REFERENCE/DOCKET NUMBER: A30042 - 165/30555

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-408-2628

TELEFAX: 212-765-2519

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2713 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE:

ORIGINAL SOURCE:

PREPARE:

PCT-US96-01735-1

Query Match Best Local Similarity 5.2%; Score 81; DB 5; Length 2713;

Matches 51; Conservative 34; Mismatches 93; Indels 44; Gaps 8;

QY 9 MQVEDTSEYNRIVLSLPLANIMEPLRHLQTLWLNRYLAGTLTYISGLMCFYIYK 68
DB 2234 MSFWSISIF-NLAVMLNLVAFIYP-----LKGVRGGLTEPRHWSGLMTGLISLG 2283
QY 69 INVYLPK-DAIPITKAMRLQMFVAMKAMPWYTLPTVSESIERG 112
DB 2284 IVIGLPHNPGIRALIGSTIRLIFSVGSOPALFLGAFNVCNKIIFLMSFVNGCGFTTRG 2343
QY 113 WTSCFASIDEFGWILYFYIAIYLVFVEGIGIWMHRELHDIDIPLYKYLHATHIYKNT 172
DB 2344 YRAVNLVLD-----VELYHLLVLYICAMGLF-----VHVFYSLLLDLVYRESL 2390
QY 173 LSPFAGLAFHPVDGIQAVPHVIALFTVPIHTTHIGLME 214
DB 2391 LNVIAKSVTRNGRSITLTAVALILVLYFSI-----VGYLEFK 2427

RESULT 8

US-08-742-440A-6

; Sequence 6, Application US/08742440A

; Patent No. 5892014

; GENERAL INFORMATION:

; APPLICANT: Coughlin, Shaun

; APPLICANT: Ishihara, Hiroaki

; TITLE OF INVENTION: Protease Activated Receptor

; TITLE OF INVENTION: 3 and Uses thereof

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Bozicevic & Reed, LLP

; STREET: 285 Hamilton Avenue, Suite 200

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94301

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/742,440A

; FILING DATE: 30-OCT-1996

; CLASSIFICATION: 536

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Sherwood, Pamela J

; REGISTRATION NUMBER: 36,677

; REFERENCE/DOCKET NUMBER: UCAL/060PAT

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650-327-3400

; TELEFAX: 650-327-3231

; TELETYPE:

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 408 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; FEATURE:

; US-08-742-440A-6

; Query Match 5.2%; Score 80.5; DB 2; Length 408;

; Best Local Similarity 20.4%; Pred. No. 0.67;

; Matches 59; Conservative 44; Mismatches 91; Indels 95; Gaps 17;

; 15 TGFYNRIVLSHLLPANLWEPRLPHFLOTWLRNVLACTLLYFISGFLMCF-----YIY 65

; 148 YGFYTLAIADFL-----FCVT-----LPFKIAYHLNMMNVFGEVLCRATTYVF 192

; 66 YKINIVYLPKDAIPITKAMRLQMFVAMKAMPYTLTPYSESIIEGKTKCFASIDEPFG 125

; 193 Y--GNNYC--SILLACISINRYLAIVHPFTYRGIPKHTYVALVTCGLV-----W 237

; 126 IYFYVIAIYLVFVEFGIYMMHRELHDIKPKYKYLHATHIYKONTLSP-----PAGLA 180

; 238 ATVFILMLPFI-----LKOEYLVOP--DITTCNDVNTCCSSSPFOLYITISLA 286

; 181 FHPVDGILOAVPHVIALFTYPIHFTTHIGLPEWAIWTANIHDCIIGNTPVYGA----- 235

; 287 FF--GFL--IPFVLIIY-----CYAAIIRTLNAYD--HRMLWYVASLLIIV 327

; 236 -----GYHTIHHYTKHNY--GHYTIWMDMM-FGSLR--DPL 268

; 328 ITTICPAPSNIIILIHANVTYNNMDGLFTYLIADLCLDSLNSCLDPFL 376

; DB

; OY

; DB

; OY

; DB

; OY

; DB

; OY

; DB

; OY

; DB

; OY

US-08-597-236-11

; Sequence 11, Application US/08597236

; Patent No. 5733765

; GENERAL INFORMATION:

; APPLICANT: STINGELE, Francesca

; APPLICANT: MOLETT, Beat

; TITLE OF INVENTION: LACTIC BACTERIA PRODUCING

; TITLE OF INVENTION: EXOPOLYSACCHARIDES

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americans

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/597,236

; FILING DATE:

; CLASSIFICATION: 426

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: EP 95201669.9

; FILING DATE: 20-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Panucci A., Allan

; REGISTRATION NUMBER: 30256

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 360 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-597-236-11

; Query Match 5.1%; Score 80; DB 1; Length 360;

; Best Local Similarity 18.5%; Pred. No. 0.64;

; Matches 46; Conservative 35; Mismatches 58; Indels 110; Gaps 12;

; 14 ETSFNRIYVLSHLLPANLWEPRLPHFLOTWLRNVLACTLLYFISGFLMCFYIY-LKINIVY 72

; 94 EITRANSVLITLIGIPKVAEH-----YFVATPLYLSLFFYATSFNI- 137

; 73 LPRDAIPITKAMRLQMFVAM-----KAMPWY--TLTPYVSESMIERGWTGCF 117

; 138 -----SROFIAGLVLAISFALDKKMPWFILVLAIVLHNA----- 174

; 118 ASIDFGWILYFYVIAIYLVFVEFGIYMMHRELHDIKPKYKYLHATHIYKONTLSPRA 177

; 175 -----TAIVAFPYWLVKIVMDV-----KLSIF-----PTTIRA 205

; 178 GLAFHPVDGI-LOAVPH-----VIALEFVPIHFTTHIGLPEMA 215

; 206 SEFFDAIILNIFVFFPHEMYITGTFNISDQCGRHVAVLKIFILDLFTL---FLFYK 262

; 216 IWTANTHDC 224

; 263 SY-ALISEC 270

; DB

; OY

; DB

; OY

; DB

; OY

; DB

; OY

; DB

; OY

; DB

; OY

; DB

; OY

; DB

; OY

; DB

; OY

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 4, 2002, 17:05:11 ; Search time 23.17 Seconds

(without alignments)
898.342 Million cell updates/sec

Title: US-09-775-879-21

Perfect score: 1562

Sequence: 1 MAADNAVLMQFVDETSFYNR.....SLRDLLEEDDKNSFKKAE 281

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: /SIDSR/gcgdata/geneseq/geneseq/AA1980.DAT:*
2: /SIDSR/gcgdata/geneseq/geneseq/AA1981.DAT:*
3: /SIDSR/gcgdata/geneseq/geneseq/AA1982.DAT:*
4: /SIDSR/gcgdata/geneseq/geneseq/AA1983.DAT:*
5: /SIDSR/gcgdata/geneseq/geneseq/AA1984.DAT:*
6: /SIDSR/gcgdata/geneseq/geneseq/AA1985.DAT:*
7: /SIDSR/gcgdata/geneseq/geneseq/AA1986.DAT:*
8: /SIDSR/gcgdata/geneseq/geneseq/AA1987.DAT:*
9: /SIDSR/gcgdata/geneseq/geneseq/AA1988.DAT:*
10: /SIDSR/gcgdata/geneseq/geneseq/AA1989.DAT:*
11: /SIDSR/gcgdata/geneseq/geneseq/AA1990.DAT:*
12: /SIDSR/gcgdata/geneseq/geneseq/AA1991.DAT:*
13: /SIDSR/gcgdata/geneseq/geneseq/AA1992.DAT:*
14: /SIDSR/gcgdata/geneseq/geneseq/AA1993.DAT:*
15: /SIDSR/gcgdata/geneseq/geneseq/AA1994.DAT:*
16: /SIDSR/gcgdata/geneseq/geneseq/AA1995.DAT:*
17: /SIDSR/gcgdata/geneseq/geneseq/AA1996.DAT:*
18: /SIDSR/gcgdata/geneseq/geneseq/AA1997.DAT:*
19: /SIDSR/gcgdata/geneseq/geneseq/AA1998.DAT:*
20: /SIDSR/gcgdata/geneseq/geneseq/AA1999.DAT:*
21: /SIDSR/gcgdata/geneseq/geneseq/AA2000.DAT:*
22: /SIDSR/gcgdata/geneseq/geneseq/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1020	65.3	333	22	AA80924 Sterol C5 desaturase
2	771	49.4	136	21	AA828356 Arabidopsis thalia
3	716	45.8	148	21	AA818089 Eucalyptus grandis
4	606	38.8	139	21	AA828354 Arabidopsis thalia
5	566	36.2	131	21	AA828355 Arabidopsis thalia
6	330	21.1	307	21	AA856950 Human prostate can
7	172	11.0	266	21	AA816769 Arabidopsis thalia
8	167	10.7	252	21	AA841385 Arabidopsis thalia
9	167	10.7	253	21	AA841384 Arabidopsis thalia
10	164.5	10.5	177	21	AA841386 Arabidopsis thalia
11	158.5	10.1	173	21	AA816771 Arabidopsis thalia

12	158.5	10.1	183	21	AA816770 Arabidopsis thalia
13	158.5	10.1	211	21	AA824357 Arabidopsis thalia
14	158.5	10.1	220	21	AA824356 Arabidopsis thalia
15	158.5	10.1	221	21	AA824355 Arabidopsis thalia
16	148.5	9.5	293	18	AA814481 Ramp-1. Ratius ra
17	147.5	9.4	265	21	AA898034 Rice SYR2 homology
18	141.5	9.1	625	18	AA818008 Arabidopsis CER1 P
19	140.5	9.0	206	21	AA807244 Arabidopsis thalia
20	140.5	9.0	206	21	AA850821 Arabidopsis thalia
21	140.5	9.0	214	21	AA807243 Arabidopsis thalia
22	140.5	9.0	214	21	AA850820 Arabidopsis thalia
23	140.5	9.0	298	21	AA807242 Arabidopsis thalia
24	140.5	9.0	298	21	AA850819 Arabidopsis thalia
25	138.5	8.9	580	21	AA830495 Arabidopsis thalia
26	138	8.8	211	21	AA831068 Arabidopsis thalia
27	138	8.8	219	21	AA831067 Arabidopsis thalia
28	138	8.8	303	21	AA831066 Arabidopsis thalia
29	137	8.8	310	21	AA857936 Human transmembran
30	136.5	8.7	258	21	AA898033 Corn SYR2 homology
31	135.5	8.7	625	21	AA803108 Arabidopsis glossy
32	135	8.6	259	21	AA898035 Soybean SYR2 homol
33	135	8.6	264	21	AA898032 Corn SYR2 homology
34	130.5	8.4	210	21	AA818064 Pinus radiata meth
35	130.5	8.4	452	21	AA830496 Arabidopsis thalia
36	129	8.3	301	21	AA838120 Arabidopsis thalia
37	129	8.3	339	21	AA838119 Arabidopsis thalia
38	127	8.1	161	21	AA898030 Arabidopsis thalia
39	126	8.1	234	21	AA806732 Arabidopsis thalia
40	124	7.9	217	21	AA838121 Arabidopsis thalia
41	123.5	7.9	216	21	AA898037 Wheat SYR2 homology
42	123	7.9	269	21	AA898036 Wheat SYR2 homology
43	122.5	7.8	573	21	AA829545 Arabidopsis thalia
44	122.5	7.8	576	21	AA829544 Arabidopsis thalia
45	122.5	7.8	604	21	AA829543 Arabidopsis thalia

ALIGNMENTS

RESULT 1	AA80924 standard; Protein: 333 AA.
ID	AA80924
XX	XX
AC	AA80924:
XX	XX
DT	28-AUG-2001 (first entry)
XX	XX
DE	Sterol C5 desaturase long partial clone protein sequence.
XX	XX
KW	Moss: Physcomitrella patens; lipid metabolism related protein; LMPB;
KW	lipid biosynthesis; lipid modification; lipid degradation; cofactor;
KW	fatty acid transport; genetic engineering; fatty acid; enzyme; plant;
KW	microorganism; polyunsaturated fatty acid; oilseed plant; maize; wheat;
KW	biotic stress tolerance; abiotic stress tolerance; rice; oat; triticale;
KW	rice; barley; soybean; peanut; cotton; rapeseed; canola; manihot;
KW	pepper; sunflower; tagetes; potato; tobacco; eggplant; tomato; vicia;
KW	pea; alfalfa; coffee; cacao; tea; Salix; oil palm; coconut;
KW	perennial grass; forage crop.
XX	XX
OS	Physcomitrella patens.
XX	XX
PN	WO200138484-A2.
XX	XX
PD	31-MAY-2001.
XX	XX
PF	22-NOV-2000; 2000WO-EP11615.
XX	XX
PR	25-NOV-1999; 99WO-EP09108.
XX	XX
PA	(BADT) BASF PLANT SCI GMBH.
XX	XX
PI	Lerchl J, Renz A, Ehrhardt T, Reindl A, Cirpus F, Bischoff F;
PI	Frank M, Freund A, Duwenig E, Schmidt R, Reski F;

XX DR WPI; 2001-367669/38.
 XX Nucleic acids encoding lipid metabolism related proteins from
 PT Physcomitrella patens useful to produce fine chemicals in modified
 PT organisms, particularly polyunsaturated fatty acids in oilseed plants -
 XX
 PS Claim 31; Page 119; 120pp; English.
 XX
 CC The present invention describes isolated nucleic acid sequences which
 CC encode lipid metabolism related proteins (LMP). The LMP nucleic acids
 CC can be used to modify lipids and fatty acids, cofactors and enzymes in
 CC microorganisms and plants, particularly to produce polyunsaturated fatty
 CC acids, and are especially useful in oilseed plants. The nucleic acids
 CC may also confer biotic or abiotic stress tolerance, particularly to
 CC maize, wheat, rye, oat, triticale, rice, barley, soybean, peanut,
 CC cotton, rapeseed, canola, manihot, pepper, sunflower, tagetes, potato,
 CC tobacco, eggplant, tomato, vicia species, pea, alfalfa, coffee, cacao,
 CC tea, Salix species, oil palm, coconut, perennial grasses and forage
 CC crops. AAH50878 to AAH50882 represent primers used in the exemplification
 CC of the present invention. AAH50883 to AAH50968 represents LMP nucleotide
 CC sequences, and AAG80843 to AAG80928 represent LMP protein sequences,
 CC given in the present invention.
 CC
 XX Sequence 333 AA:
 SQ
 Query Match 65.3%; Score 1020; DB 22; Length 333;
 Best Local Similarity 67.7%; Pred. No. 1,1e-108;
 Matches 176; Conservative 35; Mismatches 49; Indels 0; Gaps 0;
 QY 7 YLMQVDETSFYRIYLSHLLPANLWELRPHLOTWLRNYLACTLLYFISGLWCFTYIY 66
 Db 62 ylafveetwmdylvgrplssvdsrlphltqltwlrvagmlllyfvsqglwlylvs 121
 QY 67 LKINIVLPKATPTIKAMRQVAMKAMPWYLLRPVSSMTERGKTFKFSIDDEGWI 126
 Db 122 wkgehtfragdpakepdlmqdwcmkmpvycgrlclseymlergwckctafiedgwl 181
 QY 127 LYFVYIAIYLVFVEFGIYWMHRELHDIKPLKYKLTATHNHNKONTLSPFAGLAHFPVDG 186
 Db 182 tyvglaylaylavelfgymhrelhdkrlkylhathbhlhynqntslpfaqlafhpidg 241
 QY 187 ILQAVPHVIALFTVPIHFTTHIGLLEMEAIWTANIHDCINIGNIPVWAGAGYHTHTTYK 246
 Db 242 llqacphvialflflpmhffchevllfcesgwtclnhdclgdnvgimgagfhtchhctyr 301
 QY 247 HNYGHTYIMDMKMGSLRDP 266
 Db 302 hnryhylvcmwllfgllrtp 321
 RESULT 2
 AAG28356
 ID AAG28356 standard; Protein; 136 AA.
 XX
 AC AAG28356;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 33543.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX
 OS Arabidopsis thaliana.
 XX
 PN EP1033405-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131479.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137528.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138647.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.

```

PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145192.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145226.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 27-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.

```

```

PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 49.4%; Score 771; DB 21; Length 136;
 Best Local Similarity 100.0%; Pred. No. 1.3e-80;
 Matches 136; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 146 MREELHDIKPLKYLYLATHNHYKNTLSPPAGLAFHPYDGLIQAAVPIVATLPIVPIHT 205
DB 1 mhrelhdikplylylathnhykntlsppaglafhpydgliaqavpivatlpviphft 60
QY 206 THIGLLEMEALWTANTHDCIHGNIWPGAGYTHHTTKKHYGHTYIMDMMEGSLRD 265
DB 61 thigllfmealwtantahdcihgnlwpvgagylthhttkkhygnytwmdmefgsld 120
QY 266 PLEEDDNKDSFKAE 281
DB 121 plleeddnksfkake 136

```

```

RESULT 3
AAB18089
ID AAB18089 standard; Protein; 148 AA.
XX
AC AAB18089;
XX
DT 08-NOV-2000 (first entry)
XX
DE Eucalyptus grandis sterol desaturase protein SEQ ID NO:250.
XX
KW Eucalyptus grandis; Pinus radiata; modification; isoprenoid; plant;
KW metabolism; isoprenoid biosynthetic pathway; terpenoid; steroid;
KW genome mapping; physical mapping; positional cloning; forestry;
KW agriculture; medicine; fermentation; plant development; pest resistance;
KW pinene; myrcene; Monterey pine.
XX
OS Eucalyptus grandis.
XX
PN WO200036081-A2.
XX

```


PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141844.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155569.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 38.8%; Score 606; DB 21; Length 139;
Best Local Similarity 93.5%; Pred. No. 1,2e-61;
Matches 115; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
OY 1 MADNNAYLMQFVDETSFYRRIYVLSHLLPANLMEPLPHFIQTWLRNYIAGTLTYFISGFILM 60
|||||
Db 1 meadnaylmqfvdetsfyrrivlsllpanlweplphfiqtwlrnyjagtllyfissgfiw 60
OY 61 CFYIYLYLKTINYLPKDAITPTIKAMRLQMFVANKAMPWYTLPTVSESMTERGWTCCFASI 120
|||||
Db 61 cfyiylylktinylpkdaitptikamrlqmfvankampwylftptvse:imterxwxcffasi 120
OY 121 DEF 123
||
Db 121 get 123

RESULT 5
AAG28355
ID AAG28355 standard; Protein; 131 AA.
XX AAG28355;
AC
XX
XX
DT 17-OCT-2000 (first entry)
XX
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33542.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155186.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0158294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 36.2%; Score 566; DB 21; Length 131;
 Best Local Similarity 93.0%; Pred. No. 4, 5e-57;
 Matches 107; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 9 MOFVETSEFNRIIVLSHLPLANLMEPLPHFLQTWLRNLTAGTLTYFISGFLMCFYIYLYK 68
 DB 1 mqvfdetstfynivlshllpanlweplphflqtwlrnltagtltyfisgflmcfyiylyk 60
 QY 69 INVIYLPKDAIPITKAMRLQMFVAMKAMPWYTLPLTVSESMIERGWTGCFASIDEF 123
 DB 61 InvviyLPkdaipITkAmRLqmIqmfVamKamhgtlflTVseSmierGwtGcfasIdEf 115

RESULT 6
 AAB56950
 ID AAB56950 standard; Protein; 307 AA.
 AC AAB56950;
 XX
 DT 13-MAR-2001 (first entry)
 XX Human prostate cancer antigen protein sequence SEQ ID NO:1528.
 DE Human prostate cancer antigen protein sequence SEQ ID NO:1528.
 XX
 KW Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
 KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
 KW vulnereary; gastrointestinal; nephrotropic; antinefective; gynaecological;
 KW antibacterial; gene therapy; neutral; immune; reproductive; renal;
 KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
 KW wound; infectious disease.

XX Homo sapiens.
 OS
 PN WO20005174-A1.
 XX
 XX 21-SEP-2000.
 PD
 XX 08-MAR-2000; 2000MO-US05988.
 XX
 PR 12-MAR-1999; 99US-0124270.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX
 PI Rosen CA, Ruben SM;
 DR WPI: 2000-587513/55.
 DR N-PSDB; AAF16153.
 XX
 PT Prostate cancer associated gene sequences, referred to as prostate
 PT cancer antigens, useful for treatment, prevention, and diagnosis of
 PT disorders such as prostate cancer -
 XX
 PS Claim 11; Page 1969-1970; 2338pp; English.
 XX
 CC AAF15566 to AAF16505 encode the human prostate cancer associated
 CC proteins, called prostate cancer antigens, given in AAB56163 to AAB57302.
 CC The prostate cancer antigens can have neuroprotective, cytosolic,
 CC cardioactive, immunomodulatory, muscular, vulnereary, gastrointestinal,
 CC nephrotropic, antinefective, gynaecological and antibacterial activities,
 CC and can be used in gene therapy. The prostate cancer antigen
 CC polynucleotides may be used for detection of prostate cancer, chromosome
 CC identification, as chromosome markers, and for numerous other diagnostic
 CC or research purposes. The prostate cancer antigens may be used to treat
 CC disorders such as neural, immune, muscular, reproductive,
 CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
 CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
 CC AAB57303 represent sequences used in the exemplification of the present
 CC invention.
 XX
 SQ Sequence 307 AA;

Query Match 21.1%; Score 330; DB 21; Length 307;
 Best Local Similarity 32.9%; Pred. No. 2e-29; 73; Indels 42; Gaps 12;
 Matches 82; Conservative 52; Mismatches 29;

QY 51 LLYFISGFLMCFYIY-----YLKINVYLPKDAIPITKAMRLQMFVAMKAMPWYTL 101
 DB 45 llyfcaetlsyfvfdhalmkhpqflkngy-----rrel:ftvgalpwisll 91
 QY 102 PYVSESMIE-RGWTGCFASIDEGWILYFYNAI-YLVFVEGCIYWHRELHIDPLVK 158
 DB 92 -lvalilslrYsklhddlgefyglfelvvalisflftdmfiyw:hnglhh-ilyvk 149
 QY 159 YLHATHTIYNKQWTLSPFAGLAFHPVDGIQANPHVIALFTVPIHPT:HTGLLFMEAIWT 218
 DB 150 rllkphniw---ktpcfashahpIdglfIqsiPynhlypflfIphhvwYsilylvniw 206
 QY 219 ANIHDCIHGN-----IWP-VMGAGYHTHTHTYKKNYGHY-TIWMIDMEGSLRDLLEE 270
 DB 207 lsihd---gdfirvpqllqglfingsahhtdmhmfdynygqfclw-d:lgsgfkpsste 262
 QY 271 DDNKDSFKK 279
 DB 263 gKpIlsYvk 271

RESULT 7
 AAG16769
 ID AAG16769 standard; Protein; 266 AA.
 AC AAG16769;

XX 17-OCT-2000 (first entry)
DT Arabidopsis thaliana protein fragment SEQ ID NO: 17541.
XX
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135553.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.

PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-015106.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match	11.0%	Score 172	DB 21	Length 266
Best Local Similarity	22.8%	Pred. No. 2,3e11		
Matches 65	Conservative 36	Midates 84	Indels 100	Gaps 14
QY	32	WEPLPRFLQFWLR---	YVLGTLTLYFTSGFIMC-----	62
Db	9	wqylvthtsdfqlacisgflfhesvffisgdrpfiflerqgfslykqicgmtnrpaqykc	68	
QY	63	---YIYLKINVYLPKRAIPRIKAMKRLQMFAMAMAPVYTLTLPLPVSESMERGWTCKCA	118	
Db	69	ltrlllylfsvnlrplmasayvfiatmgm-----ssfp-----lps-----wkevasa	110	
QY	119	SIDEGFWLVEUYIAIVLVEFGYIWMHRELHDRIKRYLQAHATNIYNKONTSPFA-	177	
Db	111	q-----llfyfiedfeywghrhlhs-kwlykvnshvhey-----atpfgl	152	
QY	178	-----GLANPRDGLQANRVNYALFIVRIHNTIINIGLLFMEALWTANIHDCIH	226	
Db	153	tseyahpaeellfigatctypralltqrhlftwl-----wmvltvleevahcygh	202	
QY	227	-----GNIMPVN-GAGVYHTIHNT-----TYKKNNGHYTIIMDMWFGS	262	

Db	203	fwpslnftlpygsdftndyhrlllytksngysstfymdwifgt	247
RESULT	8		
AAAG41385			
ID	AAAG41385	standard; Protein; 252 AA.	
XX			
AC	AAAG41385;		
XX			
DT	18-OCT-2000	(first entry)	
XX			
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 51481.	
XX			
KW	Protein identification; signal transduction pathway; metabolic pathway;		
KW	hybridisation assay; genetic mapping; gene expression control; promoter		
KW	termination sequence.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000;	2000EP-0301439.	
XX			
PR	25-FEB-1999;	99US-0121825.	
PR	05-MAR-1999;	99US-0121810.	
PR	09-MAR-1999;	99US-0123548.	
PR	23-MAR-1999;	99US-0125788.	
PR	25-MAR-1999;	99US-0126264.	
PR	29-MAR-1999;	99US-0126785.	
PR	01-APR-1999;	99US-0127462.	
PR	06-APR-1999;	99US-0128234.	
PR	08-APR-1999;	99US-0128714.	
PR	16-APR-1999;	99US-0129845.	
PR	19-APR-1999;	99US-0130077.	
PR	21-APR-1999;	99US-0130449.	
PR	23-APR-1999;	99US-0130510.	
PR	28-APR-1999;	99US-0130891.	
PR	30-APR-1999;	99US-0131449.	
PR	30-APR-1999;	99US-0132048.	
PR	30-APR-1999;	99US-0132407.	
PR	04-MAY-1999;	99US-0132484.	
PR	05-MAY-1999;	99US-0132485.	
PR	06-MAY-1999;	99US-0132486.	
PR	06-MAY-1999;	99US-0132487.	
PR	07-MAY-1999;	99US-0132863.	
PR	11-MAY-1999;	99US-0134256.	
PR	14-MAY-1999;	99US-0134218.	
PR	14-MAY-1999;	99US-0134219.	
PR	14-MAY-1999;	99US-0134221.	
PR	14-MAY-1999;	99US-0134370.	
PR	18-MAY-1999;	99US-0134768.	
PR	19-MAY-1999;	99US-0134941.	
PR	20-MAY-1999;	99US-0135124.	
PR	21-MAY-1999;	99US-0135353.	
PR	24-MAY-1999;	99US-0135629.	
PR	25-MAY-1999;	99US-0136021.	
PR	27-MAY-1999;	99US-0136392.	
PR	28-MAY-1999;	99US-0136782.	
PR	01-JUN-1999;	99US-0137232.	
PR	03-JUN-1999;	99US-0137528.	
PR	04-JUN-1999;	99US-0137502.	
PR	07-JUN-1999;	99US-0137724.	
PR	08-JUN-1999;	99US-0138094.	
PR	10-JUN-1999;	99US-0138540.	
PR	10-JUN-1999;	99US-0138847.	
PR	14-JUN-1999;	99US-0139119.	
PR	16-JUN-1999;	99US-0139452.	
PR	16-JUN-1999;	99US-0139453.	
PR	17-JUN-1999;	99US-0139492.	
PR	18-JUN-1999;	99US-0139454.	

PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147312.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156559.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 10.7%; Score 167; DB 21; Length 252;
Best Local Similarity 22.2%; Pred. No. 8,1e-11;
Matches 58; Conservative 32; Mismatches 75; Indels 96; Caps 11;

QY 52 LXFISGFLMCF-----YIYLUKINIVLPKDAIPPIK 82
DB 25 vffisgpyiflertgtfisykqitkntpeaqqkciarlillyhcnvnlplmmasypvfr 84
QY 83 AMRLQMFVAMKAMPYUULLPVSSEMIERGWTKCFASIDFERGWLIFVYIAIVLVFVEFG 142
DB 85 fmgnme-----ssfp-----lps-----wkvvsagllfyflliedfv 114

PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0135629.
PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	19-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
<hr/>		
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147410.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0158293.
PR	13-OCT-1999;	99US-0158294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.

28-JUL-1999;	99US-0145951.
PR 02-AUG-1999;	99US-0146386.
PR 02-AUG-1999;	99US-0146388.
PR 02-AUG-1999;	99US-0146389.
PR 03-AUG-1999;	99US-0147038.
PR 04-AUG-1999;	99US-0147204.
PR 04-AUG-1999;	99US-0147302.
PR 05-AUG-1999;	99US-0147192.
PR 05-AUG-1999;	99US-0147260.
PR 06-AUG-1999;	99US-0147303.
PR 06-AUG-1999;	99US-0147416.
PR 09-AUG-1999;	99US-0147493.
PR 09-AUG-1999;	99US-0147935.
PR 10-AUG-1999;	99US-0148171.
PR 11-AUG-1999;	99US-0148341.
PR 12-AUG-1999;	99US-0148565.
PR 13-AUG-1999;	99US-0148684.
PR 13-AUG-1999;	99US-0149368.
PR 16-AUG-1999;	99US-0149175.
PR 17-AUG-1999;	99US-0149426.
PR 18-AUG-1999;	99US-0149722.
PR 20-AUG-1999;	99US-0149723.
PR 20-AUG-1999;	99US-0149929.
PR 23-AUG-1999;	99US-0149902.
PR 23-AUG-1999;	99US-0149930.
PR 25-AUG-1999;	99US-0150566.
PR 26-AUG-1999;	99US-0150884.
PR 27-AUG-1999;	99US-0151065.
PR 27-AUG-1999;	99US-0151066.
PR 27-AUG-1999;	99US-0151080.
PR 30-AUG-1999;	99US-0151303.
PR 31-AUG-1999;	99US-0151438.
PR 01-SEP-1999;	99US-0151930.
PR 07-SEP-1999;	99US-0152353.
PR 10-SEP-1999;	99US-0153070.
PR 13-SEP-1999;	99US-0153758.
PR 15-SEP-1999;	99US-0154018.
PR 16-SEP-1999;	99US-0154039.
PR 20-SEP-1999;	99US-0154779.
PR 22-SEP-1999;	99US-0155139.
PR 23-SEP-1999;	99US-0155486.
PR 24-SEP-1999;	99US-0155659.
PR 28-SEP-1999;	99US-0156458.
PR 29-SEP-1999;	99US-0156596.
PR 04-OCT-1999;	99US-0157117.
PR 05-OCT-1999;	99US-0157753.
PR 06-OCT-1999;	99US-0157865.
PR 07-OCT-1999;	99US-0158029.
PR 08-OCT-1999;	99US-0158232.
PR 12-OCT-1999;	99US-0158369.
PR 13-OCT-1999;	99US-0159293.
PR 13-OCT-1999;	99US-0159294.
PR 13-OCT-1999;	99US-0159295.
PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159337.
PR 14-OCT-1999;	99US-0159637.
PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.
PR 28-OCT-1999;	99US-0161920.
PR 28-OCT-1999;	99US-0161993.
PR 29-OCT-1999;	99US-0162142.
PR 29-OCT-1999;	99US-0162143.
PR 29-OCT-1999;	99US-0162144.
PR 29-OCT-1999;	99US-0162145.
PR 29-OCT-1999;	99US-0162146.
PR 29-OCT-1999;	99US-0162147.
PR 29-OCT-1999;	99US-0162148.
PR 29-OCT-1999;	99US-0162149.
PR 29-OCT-1999;	99US-0162150.
PR 29-OCT-1999;	99US-0162151.
PR 29-OCT-1999;	99US-0162152.
PR 29-OCT-1999;	99US-0162153.
PR 29-OCT-1999;	99US-0162154.
PR 29-OCT-1999;	99US-0162155.
PR 29-OCT-1999;	99US-0162156.
PR 29-OCT-1999;	99US-0162157.
PR 29-OCT-1999;	99US-0162158.
PR 29-OCT-1999;	99US-0162159.
PR 29-OCT-1999;	99US-0162160.
PR 29-OCT-1999;	99US-0162161.
PR 29-OCT-1999;	99US-0162162.
PR 29-OCT-1999;	99US-0162163.
PR 29-OCT-1999;	99US-0162164.
PR 29-OCT-1999;	99US-0162165.
PR 29-OCT-1999;	99US-0162166.
PR 29-OCT-1999;	99US-0162167.
PR 29-OCT-1999;	99US-0162168.
PR 29-OCT-1999;	99US-0162169.
PR 29-OCT-1999;	99US-0162170.
PR 29-OCT-1999;	99US-0162171.
PR 29-OCT-1999;	99US-0162172.
PR 29-OCT-1999;	99US-0162173.
PR 29-OCT-1999;	99US-0162174.
PR 29-OCT-1999;	99US-0162175.
PR 29-OCT-1999;	99US-0162176.
PR 29-OCT-1999;	99US-0162177.
PR 29-OCT-1999;	99US-0162178.
PR 29-OCT-1999;	99US-0162179.
PR 29-OCT-1999;	99US-0162

PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138340.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142050.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-01455913.
PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147028.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152383.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.

Query Match	10.1%	Score 158.5	DB 21	Length 183
Best Local Similarity	28.3%	Pred. No. 4.9e-10		
Matches 45	Conservative 18	Mismatches 59	Indels 37	Gaps 7
PR 26-OCT-1999;	99US-0161359.			
PR 26-OCT-1999;	99US-0161360.			
PR 26-OCT-1999;	99US-0161361.			
PR 28-OCT-1999;	99US-0161920.			
PR 28-OCT-1999;	99US-0161992.			
PR 28-OCT-1999;	99US-0161993.			
PR 29-OCT-1999;	99US-0162142.			
OY 125 WILVTVTAIYVFEVFCGYMMHRELNDIKPKYKYLHATHTHYKONTLSPPA-----	177			
DB 22 kveasqallfyfiledfvfygwhrilhs-kwlyknvshvhey-----atpfglteyeh	75			
OY 178 -----GLAFHPRDGLQAVPHVIALEFVPRHNTHTHIGLFFMAIWTANTHDCIH-----G	227			
DB 76 paeilffidatcvgpalcgrhlitwl-----mmvltvleatveahcgyhfpwsls	125			
OY 228 NIMPMV-GAGYHTHTHT--TYKNHYGHTYIMDMWFGS	262			
DB 126 nflprlyggadfhndyhnrllylksqgnysctfymdmwifgl	164			
RESULT 13				
ID AAG24357				
AC AAG24357				
DE 17-OCT-2000 (first entry)				
XX Arabidopsis thaliana protein fragment seq ID NO: 27999.				
XX Protein identification; signal transduction pathway; metabolic pathway;				
KW hybridisation assay; genetic mapping; gene expression control; promoter;				
KV termination sequence.				
XX Arabidopsis thaliana.				
OS EPI033405-A2.				
PM 06-SEP-2000.				
XX 25-FEB-2000; 2000EP-0301439.				
XX 25-FEB-1999;	99US-0121825.			
PR 05-MAR-1999;	99US-0123180.			
PR 09-MAR-1999;	99US-0123548.			
PR 23-MAR-1999;	99US-0125788.			
PR 25-MAR-1999;	99US-0126264.			
PR 29-MAR-1999;	99US-0126785.			
PR 01-APR-1999;	99US-0127462.			
PR 06-APR-1999;	99US-0128234.			
PR 08-APR-1999;	99US-0128714.			
PR 16-APR-1999;	99US-0129845.			
PR 19-APR-1999;	99US-0130077.			
PR 21-APR-1999;	99US-0130449.			
PR 23-APR-1999;	99US-0130510.			
PR 28-APR-1999;	99US-0131449.			
PR 30-APR-1999;	99US-0132048.			
PR 30-APR-1999;	99US-0132407.			
PR 04-MAY-1999;	99US-0132484.			
PR 05-MAY-1999;	99US-0132485.			
PR 06-MAY-1999;	99US-0132486.			
PR 07-MAY-1999;	99US-0132863.			
PR 11-MAY-1999;	99US-0134256.			
PR 14-MAY-1999;	99US-0134218.			
PR 14-MAY-1999;	99US-0134219.			
PR 14-MAY-1999;	99US-0134221.			

PR	14-MAY-1999	9905-0134570
PR	18-MAY-1999	9905-0134761
PR	19-MAY-1999	9905-0134941
PR	20-MAY-1999	9905-0135124
PR	21-MAY-1999	9905-0135553
PR	24-MAY-1999	9905-0135629
PR	25-MAY-1999	9905-0136021
PR	27-MAY-1999	9905-0136392
PR	28-MAY-1999	9905-0136782
PR	01-JUN-1999	9905-0137222
PR	03-JUN-1999	9905-0137528
PR	04-JUN-1999	9905-0137750
PR	07-JUN-1999	9905-0137742
PR	08-JUN-1999	9905-0138094
PR	10-JUN-1999	9905-0138540
PR	10-JUN-1999	9905-0138847
PR	14-JUN-1999	9905-0139119
PR	16-JUN-1999	9905-0139452
PR	16-JUN-1999	9905-0139453
PR	17-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139454
PR	18-JUN-1999	9905-0139455
PR	18-JUN-1999	9905-0139460
PR	18-JUN-1999	9905-0139461
PR	18-JUN-1999	9905-0139462
PR	18-JUN-1999	9905-0139463
PR	18-JUN-1999	9905-0139750
PR	18-JUN-1999	9905-0139763
PR	21-JUN-1999	9905-0139817
PR	22-JUN-1999	9905-0139869
PR	22-JUN-1999	9905-0140353
PR	23-JUN-1999	9905-0140354
PR	24-JUN-1999	9905-0140695
PR	28-JUN-1999	9905-0140823
PR	29-JUN-1999	9905-0140931
PR	30-JUN-1999	9905-0141287
PR	01-JUL-1999	9905-0141642
PR	02-JUL-1999	9905-0142154
PR	06-JUL-1999	9905-0142390
PR	08-JUL-1999	9905-0142803
PR	09-JUL-1999	9905-0142920
PR	12-JUL-1999	9905-0142977
PR	13-JUL-1999	9905-0143542
PR	14-JUL-1999	9905-0143624
PR	15-JUL-1999	9905-0144005
PR	15-JUL-1999	9905-0144085
PR	16-JUL-1999	9905-0144086
PR	19-JUL-1999	9905-0144325
PR	19-JUL-1999	9905-0144332
PR	19-JUL-1999	9905-0144333
PR	19-JUL-1999	9905-0144343
PR	19-JUL-1999	9905-0144343
PR	19-JUL-1999	9905-0144345
PR	20-JUL-1999	9905-0144352
PR	20-JUL-1999	9905-0144384
PR	21-JUL-1999	9905-0144814
PR	21-JUL-1999	9905-0145086
PR	21-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145087
PR	22-JUL-1999	9905-0145088
PR	22-JUL-1999	9905-0145192
PR	23-JUL-1999	9905-0145218
PR	23-JUL-1999	9905-0145224
PR	26-JUL-1999	9905-0145576
PR	27-JUL-1999	9905-0145513

PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160874.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.

PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161920;
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match		
Best Local Similarity 10.1%; Score 158.5; DB 21; Length 220;		
Matches 46; Conservative 18; Mismatches 56; Indels 49; Gaps 7.		
OY	125 WILVFVYIAIYLVEFEFGYVMHRELNDIKRPLYKYLHAATNHTNYKNONTLSPPA-----	177
DB	21 WKVASGAILLFYLLDFEDFVWGHRIIH-EKWLYKNVSVHHEY-----atrgltseyah	74
OY	178 -----GLAFHPDGIQAVPHVIALFIVEINHTTHIGLFLMEALWTANIHDCIH-----G	227
DB	75 paeIIlfIgfatlvgsaltpghrltIwI-----wmIlrvlveahcgylfwpss	124
OY	228 NIVPMWG-----AGYHTIHNH---TYKHNGYTIIMDMWMSGS	262
DB	125 ntlPIygsallmweafaysadfnhyhnllycksqngsstivymdwIgt	175
RESULT 15		
AAG24355		
ID	AAG24355 standard; Protein; 221 AA.	
AC	AAG24355;	
DT	17-OCT-2000 (first entry)	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 27997.	
KW	Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.	
OS	Arabidopsis thaliana.	
PID	EP1033405-A2.	
PD	06-SEP-2000.	
PF	25-FEB-2000; 2000EP-0301439.	
XX		
XX	25-FEB-1999;	99US-0121825.
XX	05-MAR-1999;	99US-0123180.
XX	09-MAR-1999;	99US-0123546.
XX	23-MAR-1999;	99US-0125788.
XX	25-MAR-1999;	99US-0126264.
XX	29-MAR-1999;	99US-0126785.
XX	01-APR-1999;	99US-0127462.
XX	06-APR-1999;	99US-0128234.
XX	08-APR-1999;	99US-0128714.
XX	16-APR-1999;	99US-0129845.
XX	19-APR-1999;	99US-0130077.
XX	21-APR-1999;	99US-0130449.
XX	23-APR-1999;	99US-0130510.
XX	23-APR-1999;	99US-0130891.
XX	28-APR-1999;	99US-0131449.
XX	30-APR-1999;	99US-0132048.
XX	30-APR-1999;	99US-0132407.
XX	04-MAY-1999;	99US-0132484.
XX	05-MAY-1999;	99US-0132485.
XX	06-MAY-1999;	99US-0132486.
XX	07-MAY-1999;	99US-0132863.
XX	11-MAY-1999;	99US-0134256.
XX	14-MAY-1999;	99US-0134218.

[illegible]

PR	26	JUL-1999	99S05-0145276
PR	27	JUL-1999	99S05-0145513
PR	27	JUL-1999	99S05-0145918
PR	27	JUL-1999	99S05-0145951
PR	28	JUL-1999	99S05-0146386
PR	02	AUG-1999	99S05-0146388
PR	02	AUG-1999	99S05-0146589
PR	02	AUG-1999	99S05-0146589
PR	03	AUG-1999	99S05-0147038
PR	04	AUG-1999	99S05-0147204
PR	04	AUG-1999	99S05-0147302
PR	05	AUG-1999	99S05-0147192
PR	05	AUG-1999	99S05-0147260
PR	06	AUG-1999	99S05-0147303
PR	06	AUG-1999	99S05-0147416
PR	09	AUG-1999	99S05-0147493
PR	09	AUG-1999	99S05-0147935
PR	10	AUG-1999	99S05-0148171
PR	11	AUG-1999	99S05-0148319
PR	12	AUG-1999	99S05-0148841
PR	13	AUG-1999	99S05-0148865
PR	13	AUG-1999	99S05-0148884
PR	16	AUG-1999	99S05-0149368
PR	17	AUG-1999	99S05-0149175
PR	18	AUG-1999	99S05-0149426
PR	20	AUG-1999	99S05-0149722
PR	20	AUG-1999	99S05-0149723
PR	20	AUG-1999	99S05-0149902
PR	23	AUG-1999	99S05-0149930
PR	25	AUG-1999	99S05-0150566
PR	26	AUG-1999	99S05-0150884
PR	27	AUG-1999	99S05-0151065
PR	27	AUG-1999	99S05-0151065
PR	30	AUG-1999	99S05-0151080
PR	30	AUG-1999	99S05-0151303
PR	31	AUG-1999	99S05-0151338
PR	01	SEP-1999	99S05-0151530
PR	07	SEP-1999	99S05-0152363
PR	10	SEP-1999	99S05-0153758
PR	13	SEP-1999	99S05-0153758
PR	15	SEP-1999	99S05-0154018
PR	16	SEP-1999	99S05-0154039
PR	20	SEP-1999	99S05-0154779
PR	22	SEP-1999	99S05-0155133
PR	23	SEP-1999	99S05-0155486
PR	28	SEP-1999	99S05-0155659
PR	28	SEP-1999	99S05-0156458
PR	29	SEP-1999	99S05-0156596
PR	04	OCT-1999	99S05-0157117
PR	05	OCT-1999	99S05-0157753
PR	06	OCT-1999	99S05-0157865
PR	07	OCT-1999	99S05-0158029
PR	08	OCT-1999	99S05-0158232
PR	12	OCT-1999	99S05-0158369
PR	13	OCT-1999	99S05-0159293
PR	13	OCT-1999	99S05-0159294
PR	13	OCT-1999	99S05-0159295
PR	14	OCT-1999	99S05-0159329
PR	14	OCT-1999	99S05-0159330
PR	14	OCT-1999	99S05-0159331
PR	14	OCT-1999	99S05-0159637
PR	14	OCT-1999	99S05-0159638
PR	18	OCT-1999	99S05-0159684
PR	21	OCT-1999	99S05-0160741
PR	21	OCT-1999	99S05-0160767
PR	21	OCT-1999	99S05-0160768
PR	21	OCT-1999	99S05-0160773
PR	21	OCT-1999	99S05-0160814
PR	21	OCT-1999	99S05-0160815
PR	22	OCT-1999	99S05-0160980
PR	22	OCT-1999	99S05-0160981
PR	22	OCT-1999	99S05-0160989

PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999;

Query Match	10.1%;	Score 158.5;	DB 21;	Length 221;
Best Local Similarity	26.9%;	Pred. No. 6.4e-10;		
Matches 46;	Conservative 18;	Mismatches 58;	Indels 49;	Gaps 7;

[illegible]

Search completed: April 4, 2002, 17:07:09
Job time: 118 sec

PR	26-JUL-1999	99S05-0145276
PR	27-JUL-1999	99S05-0145513
PR	27-JUL-1999	99S05-0145918
PR	27-JUL-1999	99S05-0145951
PR	28-JUL-1999	99S05-0146388
PR	02-AUG-1999	99S05-0146386
PR	02-AUG-1999	99S05-0146589
PR	02-AUG-1999	99S05-0147038
PR	03-AUG-1999	99S05-0147204
PR	04-AUG-1999	99S05-0147302
PR	04-AUG-1999	99S05-0147192
PR	05-AUG-1999	99S05-0147260
PR	05-AUG-1999	99S05-0147760
PR	06-AUG-1999	99S05-0147103
PR	06-AUG-1999	99S05-0147416
PR	09-AUG-1999	99S05-0147493
PR	09-AUG-1999	99S05-0147935
PR	10-AUG-1999	99S05-0148171
PR	11-AUG-1999	99S05-0148319
PR	12-AUG-1999	99S05-0148841
PR	13-AUG-1999	99S05-0148565
PR	13-AUG-1999	99S05-0148684
PR	16-AUG-1999	99S05-0149368
PR	17-AUG-1999	99S05-0149175
PR	18-AUG-1999	99S05-0149426
PR	20-AUG-1999	99S05-0149722
PR	20-AUG-1999	99S05-0149723
PR	20-AUG-1999	99S05-0149902
PR	20-AUG-1999	99S05-0149929
PR	23-AUG-1999	99S05-0149930
PR	25-AUG-1999	99S05-0150566
PR	26-AUG-1999	99S05-0150884
PR	27-AUG-1999	99S05-0151065
PR	27-AUG-1999	99S05-0151066
PR	27-AUG-1999	99S05-0151380
PR	30-AUG-1999	99S05-0151303
PR	31-AUG-1999	99S05-0151338
PR	01-SEP-1999	99S05-0151530
PR	07-SEP-1999	99S05-0152363
PR	10-SEP-1999	99S05-0153758
PR	13-SEP-1999	99S05-0153758
PR	15-SEP-1999	99S05-0154018
PR	16-SEP-1999	99S05-0154039
PR	20-SEP-1999	99S05-0154779
PR	22-SEP-1999	99S05-0155133
PR	23-SEP-1999	99S05-0155486
PR	28-SEP-1999	99S05-0155659
PR	28-SEP-1999	99S05-0156458
PR	29-SEP-1999	99S05-0156596
PR	04-OCT-1999	99S05-0157117
PR	05-OCT-1999	99S05-0157753
PR	06-OCT-1999	99S05-0157865
PR	07-OCT-1999	99S05-0158029
PR	08-OCT-1999	99S05-0158232
PR	12-OCT-1999	99S05-0158369
PR	13-OCT-1999	99S05-0159293
PR	13-OCT-1999	99S05-0159294
PR	13-OCT-1999	99S05-0159295
PR	14-OCT-1999	99S05-0159329
PR	14-OCT-1999	99S05-0159330
PR	14-OCT-1999	99S05-0159331
PR	14-OCT-1999	99S05-0159637
PR	14-OCT-1999	99S05-0159638
PR	18-OCT-1999	99S05-0159684
PR	21-OCT-1999	99S05-0160741
PR	21-OCT-1999	99S05-0160767
PR	21-OCT-1999	99S05-0160768
PR	21-OCT-1999	99S05-0160776
PR	21-OCT-1999	99S05-0160814
PR	21-OCT-1999	99S05-0160815
PR	22-OCT-1999	99S05-0160980
PR	22-OCT-1999	99S05-0160981
PR	22-OCT-1999	99S05-0160989

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 4, 2002, 22:35:46 ; Search time 3635.3 Seconds

(without alignments)
8572.380 Million cell updates/sec

Title: US-09-775-879-20

Perfect score: 1889

Sequence: 1 gaagatcgaatcaatca.....aaatgttgatgattcttgc 1889

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_lm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_lm:*
20: em_ov:*
21: em_or:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_sy:*
28: em_un:*
29: em_vl:*
30: em_htgo_hum:*
31: em_htgo_inv:*
32: em_htgo_rnd:*
33: em_htg_hum:*
34: em_htg_inv:*
35: em_htg_rnd:*
36: em_htg_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1889	100.0	2294	8	AF105034	AF105034 Arabidops
2	1876.4	99.3	3004	8	AF069468	AF069468 Arabidops
3	1876.4	99.3	103904	8	AFAC021640	AFAC021640 Arabidops
4	486.6	25.8	1175	8	AFPMASC5D	X90454 A.thaliana
5	269.2	14.3	138711	8	AFPO03214	AFPO03214 Oryza sat
6	186.2	9.9	1155	8	AF081794	AF081794 Nicotiana
7	178.2	9.4	1148	8	AF099969	AF099969 Nicotiana
8	164.2	8.7	1381	6	AX155063	AX155063 Sequence
9	132.2	7.0	310	11	G71681	G71681 A62988234FM
10	128.4	6.8	293	11	G71681	G71681 A62988234FM
11	60	3.2	7218	8	I66494	I66494 Sequence 14
12	57.6	3.0	840	8	CNS0180K	AL110675 Botrytis
13	57.4	3.0	696	8	CNS018D5	AL110681 Botrytis
14	56.6	3.0	1721	10	AB016248	AB016248 Mus muscu
15	56.6	3.0	127709	2	CNS05FC1	AL352983 Homo sapi
16	55.8	3.0	52359	2	AC010772	AC010772 Homo sapi
17	55.6	2.9	166052	2	AC013817	AC013817 Homo sapi
18	55.6	2.9	177008	2	AC011175	AC011175 Homo sapi
19	54.4	2.9	224448	2	PFMALP4	AL035477 Plasmodiu
20	54	2.9	114897	2	AP003624	AP003624 Oryza sat
21	53.4	2.8	83440	2	AC024285	AC024285 Homo sapi
22	53.2	2.8	131271	2	AC015927	AC015927 Homo sapi
23	53	2.8	152878	3	CXY18D10A	AL034393 Caenorhab
24	52.8	2.8	64612	2	AC079333	AC079333 Homo sapi
25	52.8	2.8	74119	2	AC036177	AC036177 Homo sapi
26	52.8	2.8	156550	2	AC015830	AC015830 Homo sapi
27	52.6	2.8	900	9	AF187981	AF187981 Homo sapi
28	52.6	2.8	2125	9	AB016247	AB016247 Homo sapi
29	52.6	2.8	2228	9	BC012333	BC012333 Homo sapi
30	52.6	2.8	2677	9	AF069469	AF069469 Homo sapi
31	52.6	2.8	90354	9	AF001124	AF001124 Homo sapi
32	52.6	2.8	143585	2	AC013349	AC013349 Homo sapi
33	52.6	2.8	178089	2	AC010929	AC010929 Homo sapi
34	52.6	2.8	178847	2	AP002959	AP002959 Homo sapi
35	52.2	2.8	14867	3	AE001398	AE001398 Plasmodiu
36	52	2.8	18115	2	AC023708	AC023708 Drosophil
37	52	2.8	265985	2	AC087226	AC087226 Mus muscu
38	51.8	2.7	1141	6	AX083744	AX083744 Sequence
39	51.8	2.7	65034	2	AC016067	AC016067 Homo sapi
40	51.8	2.7	81120	2	AC022851	AC022851 Homo sapi
41	51.6	2.7	835	11	CNS06JGB	AL401553 T7 end of
42	51.6	2.7	222068	2	AC087188	AC087188 Homo sapi
43	51.4	2.7	51014	2	AC011856	AC011856 Homo sapi
44	51.4	2.7	60565	2	AC023852	AC023852 Homo sapi
45	51.4	2.7	134867	2	AC013788	AC013788 Homo sapi

ALIGNMENTS

RESULT	1	AF105034	2294 bp	DNA	PLN	10-JUN-1999
LOCUS	AF105034					
DEFINITION	Arabidopsis thaliana delta7 sterol C-5 desaturase (STL7) gene,					
ACCESSION	AF105034					
VERSION	AF105034					
KEYWORDS	complete cds.					
SOURCE	AF105034.1 GI:5031218					
ORGANISM	thale cress.					
	Arabidopsis thaliana					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.					
REFERENCE	1 (bases 1 to 2294)					
AUTHORS	Choe,S., Noguchi,T., Fujioka,S., Takatsuto,S., Fiesler,C.P.,					
	Gregory,B.D., Ross,A.S., Tanaka,A., Yoshida,S., Txx,F.E. and					
	Feldmann,K.A.					
TITLE	The Arabidopsis dwf7/stel mutant is defective in the delta7 sterol					
JOURNAL	C-5 desaturase step leading to brassinosteroid biosynthesis					
	Plant Cell 11 (2), 207-221 (1999)					

MEDLINE 99128169
PUBMED 9927639
REFERENCE 2 (bases 1 to 2294)
AUTHORS Choe, S., Tanaka, A., Gregory, B. D. and Feldmann, K. A.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-1998) Plant Sciences, The University of Arizona,
Forbes Hall 303, Tucson, AZ 85721, USA
FEATURES
source
1. .2294
/organism="Arabidopsis thaliana"
/cultivar="Ws-2 (Masslenskija-2)"
/db_xref="taxon:3702"
/chromosome="III"
/map="2.2"
join(43. .368,736. .1041,1396. .1889)
/gene="STE1"
/product="delta7 sterol C-5 desaturase"
43.1889
gene
/gene="STE1"
/note="DWARF7"
join(143. .368,736. .1041,1396. .1709)
/gene="STE1"
/note="steroid/brassinosteroid biosynthetic enzyme"
/codon_start=1
/product="delta7 sterol C-5 desaturase"
/protein_id="AAD38120.1"
/db_xref="GI:5031219"
/translation="MAADNAYLMQFVDETSEYRIVLSHLLPANLMEPLPHLOTWLR
NVLAGFLYISGFELMCFYIYLKINVYLPRKQALPIIKAMRLQMFVAMKAMPYTLPL
TVSEMIERGKNTSPASIDEPGLILFYVATILYVEGCIYMHLEHDIRKLYKL
HATHHTYKNTLSPFAGLAFHPVDITLQVPIVIALFVPIHFTTHIGLMEALWT
ANIHDCIHNIMVPGAGYHTIHTTKYKHYHTIWMDFSLRDLPLEEDNDKS
FKKAE"
322
variation
/gene="STE1"
/note="dwf7-2 (ste1-3) allele"
/replace="a"
1552
variation
/gene="STE1"
/note="dwf7-1 (ste1-2) allele"
/replace="a"
BASE COUNT 585 a 469 c 417 g 823 t
ORIGIN
Query Match 100.0%; Score 1889; DB 8; Length 2294;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1889; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gaagatcgaatcaatcaatcaatctctctgctgacacatgacatctactgctgact 60
DB 1 GAAGATCGATCAATCAATCAAACTCTGCTGCCACATGATCACTACTGTTGACT 60
QY 61 ttttcaataaagtaagtaagtaagtaatcgcgcgaatctcttctcttccggcaccga 120
DB 61 TTTTCAATAAAGTAAGTAAGTAATCGCGCAATCTTCTTTCGTTTCCGGCACCGA 120
QY 121 tctcgtgtagctctcgaatcgaatgcgcgcgaataatgcttatctgatactgagttgtga 180
DB 121 TCTCGTGTGAGTCTCCGATTCACATGGCGGGAATATGCTTATCTGATGCAAGTTGTGGA 180
QY 181 cgaacaccttttcaacacgaatcgctctgagtcacatcttctcgcgcgaatcctatgga 240
DB 181 CGAAACCTCTTTTACAAACGGAATCGCTTCAAGTCATCTTTTCCGCGCAATCTATGGGA 240
QY 241 acccttaaccatcttctcgaacatgctcgcgaatctactgcgcgaacaccatata 300
DB 241 ACCCTTACCTCATTTTCTTCCAGACATGGCTCCGAATTTACCTCGCGCAACCTCATATA 300
QY 301 ctctacatccggttctctcgtgtagctctacatctactacttaaaatcaacgcttact 360
DB 301 CTTCATCTCGGTTTCTCTCGTGTGCTTCTACATCTATTACTTAAATCAACGTTTACT 360

QY 361 tcccaagctcgcacttcaacttctgtaactatctgtaactatctgtaactatctgtaact 420
DB 361 TCCCAAGGCTCGACTTCACTTTCTATTAATGCTTAACTGCTTCTATGTTAT 420
QY 421 cgatttcaatttaagaagaagggttctctgtagctctgtaagtaatttgatctgac 480
DB 421 CGATTTTCAATTTAAAGGAAGGCTTCTGTAAGTCTGATCAATTAATTTGATTTGAT 480
QY 481 gtgtagatctcatgttgcatttcatatctgtaactatctcctcatcctaaaggatg 540
DB 481 GTGATAGTTCATCTTTGCAATTTATTTATTTATTTGTCATATCTCCATTAAGGATTCG 540
QY 541 aacagtagtgcctatataagttcttgcgaacaaatgaagatcgtaactcttgaag 600
DB 541 AACAGTTAGTGCTTATATAGTTTGTGCAACCAATGAGAGTCTACATCTTTGAAG 600
QY 601 tgaatttctacttgccttaagtcacatnaatgcttgctgtaagtagtctgact 660
DB 601 TGAATTTTCTACTTGCATTTAAGTCCACTTAAATGTTGTAAGTAGTGTCTACT 660
QY 661 ttcagacacatcttctcgtctcctgagactcttctgaattgaactcttcttg 720
DB 661 TTCAGACATCTTTTCTGCTCTCTGAGACTCTCTTATGTTGAATCTTTTGG 720
QY 721 tctgtttgctcgaatgcaatctcacaatgaagtagtctgcaaatgcttg 780
DB 721 TCTGTTTGTTCGATGATGCAATTCCTCAATTAAGGCTATGCTTCAATGTTGTG 780
QY 781 caaagaagctatgacatgtaactcttccaaactgctcgcgaagtagatgaac 840
DB 781 CAAGAAGGCTATGCCATGATGATCTCTTCCAACTGCTCCGAAGATGATGATGAAC 840
QY 841 gtggtgaccacaaatgttctgtaagtaagtaagtaagtaagtaagtaagtaagtaag 900
DB 841 GTGTTGACCAAAATGTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 acatgcacatctatctgcttctgtagtctgtagtctgtagtctgtagtctgtagtctg 960
DB 901 ACATGCCATCTATCTGTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 960
QY 961 atgacatgaagcctctatgaatgtaactcgaatgcacacacacacacacacacacac 1020
DB 961 ATGACATTAAGCCTCTTATATGATCTCCATGCCACCATCATATTTACACACAGACA 1020
QY 1021 ataacctctcctatctgacgtaaggttctcagttgtcttcttcttcttcttcttct 1080
DB 1021 ATACACTCTCTCATTTGCGGTAAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1080
QY 1081 agatgtagatctgtagtcttctcacaagaagaactctgcaagactgcttctcaca 1140
DB 1081 AGATTGTTACATTTAATTTCTTACAGAAAGACTTTGTCAGACGCTGTTGATCTCCA 1140
QY 1141 aatcacatcttgcatactcctatcacaatgaatgaacgaagaagctatgaatgaat 1200
DB 1141 AATCACATTTTGTCTTCTTATCTCATTAAGTAACACAAAGGCTAGAAATTTATTAATGT 1200
QY 1201 cagctgcatctacatcatalgtlcaagaagaactctgacttaacgaagcttgaatct 1260
DB 1201 CAGCTGATTTACTTACATATGTCAGAGACCTTCTGATTAACCAAGTTTATGATCTTT 1260
QY 1261 gtgttctctctggtctcgcgaactgattggaatgacgaagaagcttcttctatctact 1320
DB 1261 GTGTTTCTCTTCTGCTCTCGACGATGATGAAGTAAGAAAGCTTCTTATCTACTTTC 1320
QY 1321 ctggaagtatctctggttcaatcacaagaatgtaacatcattacttgaactctct 1380
DB 1321 CTGGAGTGTATCTTGTGTTATCCAGAGATGTGACATTTAATTTACTTGAATCTCTTA 1380
QY 1381 cgcttctgcttcaagagctgcatcttcaacagtagaagcgaatctcaagctctgtaac 1440
DB 1381 CGTCTTTGTTTACAGGCTTGCATTTCAACCAAGATGATGATGATGATGATGATGATGAT 1440
QY 1441 atgtgataagcgtgttctatagtgccaatcattcacaactcatabatgcttcttctca 1500

Db 2060 CATGACATTAAAGCTCTTAAATATCTCCATGCCACCATCATATATCTACAAACAGAC 2119
Qy 1020 aatacaatctccatcattcgccggaagatgcttcaagttctctctctctctctctctctaa 1079
Db 2120 AATACACTCTCTCCATTTGCCGGTAAGTGTTCAGTTTGTCTCTCTCTCTCTCTCTCTAA 2119
Qy 1080 aagaatggtgagcatttgcttctctctctctctctctctctctctctctctctctctcc 1139
Db 2180 AAGATGGTGAAGATTGTTTCTTACAGAAACATTTCTTACACACTCTCTCTCTCTCTCTCC 2239
Qy 1140 aatcacattctccatcctaa 1199
Db 2240 AATACATTTTTCATTTCTTATCCATTAAGTAACAGAAAGCTTAATATTAATG 2299
Qy 1200 tcaagctgcatct 1259
Db 2300 TCAGCTCATTTACTTACATATGTCAGAGAGACTTCTGACTTACAGAGATTGATCTT 2359
Qy 1260 tctgtc 1319
Db 2360 TGTGTTCTCTTCTGCTCTCGAGTGTGAAATGACGAACTCTTATCTCTCTCTC 2419
Qy 1320 cctggagtgatct 1379
Db 2420 CTTGAGCTGATCTTGTGTTATCAAGAGATGTGACATCTATATTTACTTGTACTTCTT 2479
Qy 1380 acgcttc 1439
Db 2480 ACCTTTTGTGTTACAGGGCTTGATTTACCCAGTACAGGGATCTTACAGGCTGTACCG 2539
Qy 1440 catgtgataagcgtc 1499
Db 2540 CATGTGTACCGCTGTTTAAAGTCCAAATTCATTTCAACATCTATTAAGCTTTTCTTTC 2599
Qy 1500 atggaacgataatgagcggcgaacatccatcagatcagatcagatcagatcagatcagatc 1559
Db 2600 ATGGAACGATATGAGAGGAGGAGCAATCCATGATGATCCATGCAATCTGCGCACTA 2659
Qy 1560 atgggtgagagatacagatacagatacagatacagatacagatacagatacagatac 1619
Db 2660 ATGGGTGAGAGATACATACATGATACACACAGACATACAGATACATGATGATGAT 2719
Qy 1620 accataagatgagatgagatgagatgagatgagatgagatgagatgagatgagatgagat 1679
Db 2720 ACCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2779
Qy 1680 aacaagaacagatcagaagaacagagatgagaatgagatgagatgagatgagatgagat 1739
Db 2780 AACAAAGACAGCTTCAGAAAGAGAGATGAGATGAGATGAGATGAGATGAGATGAGAT 2839
Qy 1740 tctgtc 1799
Db 2840 TTGCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 2899
Qy 1800 tattcaatgct 1859
Db 2900 TATTCAATGCTCTCTCTCAACCTTTCATTAATTAATTAATTAATTAATTAATTAATTA 2959
Qy 1860 taaacaatgataatgataatgataatgataatgataatgataatgataatgataatg 1889
Db 2960 TAAACATGTAATGTTTGTATGATCTTTC 2989

RESULT 3
LOCUS ATAC021640
DEFINITION Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence.
ACCESSION AC021640
VERSION AC021640.7 GI:12408747
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

1 (bases 1 to 103904)
Lin. X., Kaul, S., Town, C. D., Beilto, M., Greasy, T. H., Haas, B., Wu, D., Bowman, C. L., White, O., Nierman, W. C., and Fraser, C. M.
Arabidopsis thaliana chromosome III BAC F16B3 genomic sequence
Unpublished
2 (bases 1 to 103904)
Lin. X. and Kaul, S.
Direct Submission
Submitted (18-JAN-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, xlinetlgr.org
3 (bases 1 to 103904)
Lin. X.
Direct Submission
Submitted (24-JAN-2001) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Jan 24, 2001 this sequence version replaced gi:12280770.
Address all correspondence to: atetlgr.org

BAC clone F16B3 is from Arabidopsis chromosome III and is near the molecular marker m174. The orientation of the sequence is from SP6 to T7 end of the BAC clone.

Genes were identified by a combination of three methods: Gene prediction programs including GAIL (available by anonymous ftp from arthur.epm.ornl.gov), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://genomic.stanford.edu/~chris/GENSCAN.html>), and NetPlantGene (<http://www.cbs.dtu.dk/netgene/cbsnetgene.html>), searches of the complete sequence against a peptide database and the Arabidopsis EST database at TIGR (<http://www.tigr.org/db/at/at.html>).

Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by more than two gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats are identified by RepeatMasker (Arjan Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>).

Location/Qualifiers

1. 103904
/organism="Arabidopsis thaliana"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="III"
/map="m174"
/clone="F16B3"
join(<129..241..586..>979)
/gene="F16B3.1"
<129..>979
/gene="F16B3.1"
/note="predicted by genefinder"
join(129..241..586..979)
/gene="F16B3.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="AF32445.1"
/db_xref="GI:6957701"
/translation="MEKDESGPLSENGAGLCYDDEYFGGPAKLOFRLSEIG
ELDLDDEDDVVISLKDIYGEIRAGGCGMNEYRRLKGLGILYLRHGVMTTKY
AVNTTPSDEDESICAAEFPODRDSYTKLSIDMHCARPFVYLPNSQFKRSSPGE
SEVTCFSG"
complement(join(<3079..3438..3582..>4265))
/gene="F16B3.2"
complement(<3079..>4265)
/gene="F16B3.2"
/note="Identical to COL2 GB:AA67879 from [Arabidopsis

CDS
complement(join(3079..3438,3582..4265))
/gene="F16B3.2"
/codon_start=1
/product="COL2"
/protein_id="AAF32446.1"
/db_xref="GI:6957702"
/translation="MLKEESNESGTWMAACDTCRSAACTVYCEADSAVLTCTCDARV
AANVARSHERVRCOSCESAPAFILCKADAAISCTGDAEIHSAANPLARHOBVPL
PLANSQSSMAPSEPTDADNDDEDEVSALLPNCKNIGNONNPFRLGVEVLDVLYS
SSMDNDFEDNQTTHYORSFGDGVYPLQVESTSHLOOSQNFQGLNGVSSGAHN
NNSLKLHNSASVSMDSVPESTASDITVQHRTTKETIDOLSGPTQVVOQLTPM
EREARVLYREKKTKREKTIIRYASRAVAEIRPRIKGRPAKRIETEAEEIETIS
LMSETGCIYPSF"
5228..5255
repeat_region
/rpt_family="(CAA)n"
/db_xref="GI:6957728"
/gene="F16B3.3"
/db_xref="GI:6957703"
/note="predicted by genefinder"
5471..7228
CDS
/gene="F16B3.3"
/note="hypothetical protein"
/codon_start=1
/protein_id="AAF32447.1"
/db_xref="GI:6957703"
/translation="MVPLPLKLDFTQGRAGDSGFKRGSTIRIGRIYRGEIAIKDA
GISTKRLIVDSNMIIHDLGSSNGTILNSDTIDSDTPVLSHGEDIKLEIVSILV
NFGSDVQAQOEHLPPRPNNRNLASDPDPPIESVQKPKRTGSSQOEKNELP
KSTPASKRKLDELADKEEELDIIEKVYKAPVRPRKNASAIKKEEVVEKKRG
KPRKNASATLEEEVEEVEEKGNSRARRKSEIVOKSIKLEVEDTPKAVIEEKKR
KRYRSKQIENECEGLEVKDEKRTTRSTRTSTRTSTRTSTRTSTRTSTRTSTRT
KMEPEKTRNDADGEVLKNCVDEKNEAEGSGSRSDICDDDEDEKCSKR
VEOYEILKRTSTYELKCTDKEDGETNSDITRERENENKEAQECSESRSEYE
RVGGARAVEQVEIELRKSTVGEDNLCTVRREDGEMENLOEIEECDESDCKVEE
AGPATLEEKVYQGMNKKVYERVEVYLEKMLREMPDAIEVOLPKQTEETEKMIEM
RSKSMRVKHIAEOKEKD"
complement(join(<7759..7920,8001..8091,8192..8335,
8796..8861,8964..9101,9182..9283,9433..9507,9638..9718,
10236..>10518))
/gene="F16B3.4"
complement(<7759..>10518)
/gene="F16B3.4"
/note="predicted by genscan, contains Coesterase domain"
complement(join(7759..7920,8001,8091,8192..8335,
8796..8861,8964..9101,9182..9283,9433..9507,9638..9718,
10236..10518))
/gene="F16B3.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="AAF32448.1"
/db_xref="GI:6957704"
/translation="MQLSPERCPRMSENREANSSEEMELLHGSNRLSPSEHYRRV
SGNSSEDSERICQOQSGFRDIGHAAETVITLISFLGLGIDLTIPTSGLPK
VVVFTVGAMITIGKAWSLGLQLAEDDITVACLDINFPQGITSDVSAAGISF
VCNNISAFGDPNRIYLMGOSAGAHISCAFPQAKISRESISWMSQIKAYGLS
GCYNLFWLHEHNRGLXRSIFLISIMGESEFPKPSRVRRLAKAALDELPHILL
PHGSADYSIPEASKRTTDLQAAEVAKELVMTGKTHITDPLDPLNKGADLDELHDI
VSMIHADSDALRDAAVAPPRKLVPEFLKLKLAQVSPF"
complement(join(11800..11972,12355..12394,12527..12640,
12750..12834,12919..13039,13133..13258,13559..13659,
13745..13874,14237..14456))
/gene="F16B3.5"
complement(11800..14456)
/gene="F16B3.5"
complement(join(11820..11972,12355..12394,12527..12640,
12750..12834,12919..13039,13133..13258,13559..13659,
13745..13874,14237..14413))
/gene="F16B3.5"
/note="unknown protein"
/codon_start=1
/protein_id="AAF32449.1"

DB
/db_xref="GI:6957705"
/translation="WAEGBEDSORLKIITAAAFDVENJARMADYWSNLIIPHHASRP
EVDFKRRFYORYIDPDLVYPMNSTSSQSPAPRTFTSASTASNNANOVRRNS
GSVPRTSGPSATGATPPSSMRDEQTOFVSNAWFIIVAVLPLIPKRNINRATRL
SFMGTACSSLSVLSYGRPRAMNMOGLQVYFQSIYAKDFIYICLTFVSHLCLK
FALPIILICRALEQVAKFIRNNGRSTIRKYRLEDECVAVESNTTILNLSQAETAG
FLIISLSMQRNIITOFEMYMQLIKLMYQAPVTAQYHJSTMIRIGRTVTPILQVAPF
LNTPSAVORMMER"
15001..15660
mRNA
/gene="F16B3.6"
/db_xref="GI:6957706"
/note="predicted by genefinder"
15001..15660
CDS
/gene="F16B3.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="AAF32450.1"
/db_xref="GI:6957706"
/translation="MSALRLRNANPAPLELSDQTSSESRVLKROMSQRAMNT
LTSANLISNLIPGTLIAFOLLTPYFTSNGVCDHATRTITAVILFLAASCTVSTFD
SVKADGTITRGFTVFKGMVYDIPDSGLIPLDLAKIRMRVDMITHATLSLVFGAV
ALRDKYITDCFTPEPEAEKTHVLIDIVPGVGVCSLLMVFPARRHGIGYLVTSVDR
R"
complement(join(<16445..16464,16657..17107,17220..>17870))
/gene="F16B3.7"
complement(<16445..>17870)
/gene="F16B3.7"
/note="predicted by genscan"
complement(join(16445..16464,16657..17107,17220..17870))
/gene="F16B3.7"
/note="hypothetical protein"
/codon_start=1
/protein_id="AAF32451.1"
/db_xref="GI:6957707"
/translation="MELPATKRLIGLVIFPILLITIAVILYLFPGYPLYSTYKHL
SNSSSSAISSPSRNHSSSDSKTDESPSSDNIYDDYHDPKSSLHNNRLS
ISSNGHOVTPKKEKRRKRRKCDIFSQEIPIPKIPIYNTICRAIHONCIKY
GRDLPKMKRMPKRECDLPDPPEPLIYVGTGMAIYGVDSVNNHOSLCLISRV
EHPEDSQOERFQWKYKTYNTITATFTWTHLVAAEITETGPTGNSYNIYLDLSDP
PTWASQIGEPDIILISSQWEPRLPLEDKORICGLCYIPGVNRVGAHFVRRALR
TTFKTIILLENKFGVFLPTEAPSHFEDEGEMDKAMRIS"
complement(17673..17778)
/rpt_family="(CAT)n"
repeat_region
complement(18048..18074)
/rpt_family="(TA)n"
Query Match 99.3%; Score 1876.4; DB 8; Length 103904;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1888; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 gaagatcgatcaatcaatcaatcaatcaatctctgtgtgacacatgacatctactgttgact 60
|||||
DB 62547 GAAGATCGATCAATCAATCAATCAATCTCTGTGTCCACATGCATTAATACGTGTGACT 62606
QY 61 tttcaataaagtaagaagaatcaatccgagcaatctctctgttcttcgagcaccga 120
|||||
DB 62607 TGTTCATATAAGGTAAAGTAAAGTCAATCGCGGAATCTTCTTCTTCCGGACCGA 62666
QY 121 tctcgttgatctccgattcaatcagtcgagcgcggaataatgtctatctgagt cagttgttga 180
|||||
DB 62667 TCTCGGTGATCTCCGATTCACATGCGGGGGAFTAAATGCTTATCTGANTC CAGTTGTGTA 62726
QY 181 cgaacctcttttaacaccgaatcgctctgagtcactctttgcgagccgaatctatggga 240
CGAAACCTCTTTTACAAACGAATCTTCTGACTCAATCTTTTCCGCGCAAAATATATGCGGA 62786
QY 241 acccttaccatcttctccagacatgagtcgcgaataatccctgcgaggaacctactata 300
|||||
DB 62787 ACCCTTACCTCATTTTCTCCAGACATGAGCTCCGAAATTTACCTTGCCGGAACCTTACTATA 62846
QY 301 cttcatctcgggttctcctgtgtctctacatctatctaacctaaatcaacgttacct 360
|||||

Dh	62847	CTCATCTCCGGTTTCCTCTGGTGGTTCACATCTAATACCTTAATCAACAGCTTACCT	62906
Qy	351	ttcccaaaagctctcgacttcaacttttgtaatcaccatgcttaacgcgtttctatglat	420
Dh	62907	TCCCAAAAGGTCGCACTTTCACATTTTGATTCACATATGCTTAAACGGCTTCTATGTTAT	62966
Qy	421	cgattcttccaatt-aagaaagaggctctctcgtaactgtaacagataattggaattga	479
Dh	62967	CGATTTTCANTTAAAGGAAGGGTTTCTCGGTACTGTACGTATTTGGATTTGA	63026
Qy	480	ctgtgaatgltcaatglttgacatlaatgataattgtgcatactccactcaaggatc	539
Dh	63027	TGTGATATGTTTCATCTTTGGCATTTTATGATTAATTTGTGCAATTCCTCATCTAAAGGATTT	63086
Qy	540	gaacgcttagtgagcttatataagattttgtgaaaccaaagaaagtcgtaactttgaa	599
Dh	63087	GAACGTTAATGCGCTTATATATAGTTTGTGCAACCAATGAAGAAAGTCGACATCTTTGAA	63146
Qy	600	gttgaatttctactgtcgacttaagtcgaacttaacttgattgtgtgaagtgaatgtctaac	659
Dh	63147	GTTGAATTTTCTACTTGGCATTTAATGATTAAGTCACATTAATTTGTTTGAAGTGTGCTAC	63206
Qy	660	tttcgaacaacttcttttctgtctctctcgagacctgtcttagtttgaactcttttg	719
Dh	63207	TTTCGACACATCTCTTTTCTGCTTCTGTGAACACTGTGCTTAATGTTAAACTTTTGTG	63266
Qy	720	gtctatcttgctctcagatgagcaatctccataaagagctatgcgttttgcgaattgtctgt	779
Dh	63267	GTCGTCTTTGCTTCAGATGACATTTCTCTCAATAAAGGCGATCCGTTTGCMAATGTTTGG	63326
Qy	780	gcaatgaaagctatgcaatgltglaacactcttctccaactgtctccagagatgatatgaa	839
Dh	63327	GCAATGAAGGCTATCCACTGTGACACTCTTCTTCCAACTGTCTCCAGAGTATGATTTGAA	63386
Qy	840	ctgtgtctgaccaaatgatttctgtagaaagaacacacacacacacacacacacacacacac	899
Dh	63387	CGTGGTTGGACCAATGTTTTCCTAGCATAGGCGAATTCGGTGGATTCGTATTTTGT	63446
Qy	900	tcaatcgcaactaactctgttcttgatgagtttgaattatlaattgaaatgcaagagactc	959
Dh	63447	TACATCGGCATCTACTTGTGTTTCCCTTAGTTGGTATTAATTTAGGATGCACAGAGAGCTT	63506
Qy	960	catgactttaaagccctctaaagaatctccatgcaacccatcatatctaaacaaagag	1019
Dh	63507	CATGCACTTAAAGCCCTCTATATAGATCTCCATGCGCACCCATCATATCTACACAAAGCG	63566
Qy	1020	aataaactctccaatttgccggtgaagtgttctagttgttctctttagtctctgttaa	1079
Dh	63567	AATACACTCTCCATTTGCGCGTAAGTGTTCAGTTGTTCTTCTTATGTTCTTGTA	63626
Qy	1080	aagatctgttagaatttagtttctcttaacgaagaaagacttgttaagaactgtcttgaacc	1139
Dh	63627	AAGATTTGATGATTTATGTTTCTTAACCGAAAGAAAGCTTTGTAGAGAGCTGTATCTCC	63686
Qy	1140	aaatcacacttctgcaatctccatlaacaaagaatcaacagaagctcagataatacaaatg	1199
Dh	63687	AAATACACTTTTGCATTTCTCTTATCCATATAAGTAACCAAGAAAGCTATGATTTATTAATG	63746
Qy	1200	tcaactgcataactcacacatgltcagaagagactctcgacttaaccagagtttagactct	1259
Dh	63747	TCAAGTCGATTACTTCAATATATGTCAGAGAGACTTCTGACTTAAACAGATTTAAATCTT	63806
Qy	1260	ctgattcttctctgtgctctgagacgattggaatggaagagaagtcctttaaactacttc	1319
Dh	63807	TGTGTTTCTCTTCTGTCTCGGACTGATTTGGAAATGACGAGAGATTCCTTTATCTACTTC	63866
Qy	1320	ctgtgagatglaatctgtgtlaatccaaagatgagacaactaataacttgttaactctct	1379
Dh	63867	CCTGAGATGATCTTGGTTAATCAAGAGATGACATCTAATTAATTAATCTTAATCTCTT	63926
Qy	1380	acgttcttggttaacaggctgcacttcaaccagtagacgggaatcttcaaggctgtacgg	1439
Dh	63927	ACGTTTGTGTTTACAGGGCTTGCATTTTACCCAGATGAGCGGATCTTTCAGGCTGTACCG	63986

OY	1440	caagtgatagcgcgttttataatgccaattcaattcaaacctatagaagctcttgctc	1499
Db	63987	CATGTCGATAGCCCTTTTATATAGTGCCTAATTCATTTTCACACTCATATATAGCTCTTTGTTTC	64046
OY	1500	atggaagcgatataggacaggcgacacatccatgactgcataccatgtagcaacatcctggccagta	1559
Db	64047	ATGGAAGCGATATGGAACGGCGAACAATCCATGACTGATCCATGCAATGGCAACATCTGGCCAGTA	64106
OY	1560	atggatgcagaatlaaccaaacgataaaccaagaacatataagacataactatggtcaattat	1619
Db	64107	ATGGGTGGAAGATACCAATACGATATACCAACAAGAAATACAAACATATACATATGATATAT	64166
OY	1620	accatcatgatagcatcttgatgtcttgagccctcttaaggatccctctctgaagaagaatgac	1679
Db	64167	ACCATATGATGATGATGGATGCTTTGGCTCTCTTAGAGATCCCTCTTGAAGAAAGATGAC	64226
OY	1680	aacaaagacagcttcaagaaagcagatgtagaaatgacccactggatcttgctctctcgtt	1739
Db	64227	AACAAGACAGCTTCAAAAGAAACAGATGAGATGCCCTTGGGGTTTGTCTTCTGT	64286
OY	1740	ttgtctcttgatgtctgttgcataaagtttcaagccttcttgctctctctctctctct	1799
Db	64287	TGTGTTGTTGTTGTTGTTGTTTAAAGTTTCACCTTTCTGTGCTTTTCTCTCTCTT	64346
OY	1800	tattcatatgtctctctcctaacccttccaatatataatgttatacaaaatctgctcgtctagtt	1859
Db	64347	TATTCATGATGCTCTCTCAACTTTCCAATATATATTTGATTAACAACATTTGCTGTCTAGTT	64406
OY	1860	taaaacatgtaaatgattcttgatgatactcttggc	1889
Db	64407	TAAACATGTAATAATGTTGATGATCTTTTGGC	64436

RESULT	4
LOCUS	ATRNASCSD
DEFINITION	ATRNASCSD 1175 bp mRNA
ACCESSION	A.thaliana mRNA for sterol-C5-desaturase.
VERSION	X90454
KEYWORDS	X90454.1 GI:1061037
SOURCE	sterol-C5-desaturase gene.
ORGANISM	thale cress.
REFERENCE	Arabidopsis thaliana
AUTHORS	Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Rosidae: eurosids II; Brassicales; Brassicaceae; Arabidopsis.
TITLE	1 (bases 1 to 1175)
JOURNAL	Gachotte,D., Husselstein,T., Bad,M., Lacroute,F. and Benveniste,P.
MEDLINE	Isolation and characterization of an Arabidopsis thaliana cDNA encoding a delta 7-sterol-C-5-desaturase by functional complementation of a defective yeast mutant
REFERENCE	Plant J. 9 (3), 391-398 (1996)
AUTHORS	97077348
TITLE	2 (bases 1 to 1175)
JOURNAL	Gachotte,D.
MEDLINE	Direct Submission
REFERENCE	Submitted (12-JUL-1995) D. Gachotte, Institut de Biologie Moléculaire Plantes, Enzymologie Moléculaire et Cellulaire, Institut de botanique 28, rue Goethe, F- 67083 Strasbourg, FRANCE
FEATURES	Location/Qualifiers
SOURCE	1. 1175
	/organism="Arabidopsis thaliana"
	/variety="Landsberg erecta"
	/db_xref="taxon:3702"
	118..963
	/codon_start=1
	/product="sterol-C5-desaturase"
	/protein_id="CA62079.1"
	/db_xref="GI:1061038"
	/db_xref="SPTREMBL:Q39208"
	/translation="MADNAYLWQFVDENSFYKRIYVLSHLLPANIWEPLPHIQTLNR
	NYAGTLIYISGLFWCEYIYKRIINVEYKRDIPITKARLQMFVAMKAMPYTLIP
	TVTSMERWMTKCFASIGGEFGLIIFVYILVAVELVEPIYMMHEHLHDIKPLVYV

HATHHIYKONTLSPFAGLAFHPVDGILQAVPHVIALFIVPHFTTHIGLLEMEIWA
ANIHDCIGHNIPWGMGAGYHTIHHTYKHNHYGHTIWMOWFSLRDLLEEDNKRKS
FKKAE"

BASE COUNT 300 a 265 c 227 g 383 t
ORIGIN

Query Match 25.8%; Score 486.6; DB 8; Length 1175;
Best Local Similarity 98.2%; Pred. No. 2.3e-93;
Matches 492; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1389 tttaacagagcttgatcattccaccagtagacgagatcttcagctgctacccagatgta 1448
Db 643 TTTCGGGGCTTGATTCACCCAGTAGACGGATACTTGACGGCTGATCCCATGTGATA 702
OY 1449 ggcgtgtttatagtgccaattcattcaacaatcataataggtcttcttgatggaacg 1508
Db 703 GCGCTGTTTATAGGCCAATTTCATTCACACTATATAGGCTTTTCTTCATGGAACCG 762
OY 1509 atatgagcgcgaacatccatctgactgcatcagcaacatctgcccagtaatggtgca 1568
Db 763 ATATGGGCGCGAATCATCTGATGCTCATGCGAATCTGGCCATTAATGGGTGCA 822
OY 1569 ggaatacctacgatacaccacagacacacataactatggtcattatataccatag 1628
Db 823 GGATACCAATGAGATACACACACACATACAGCATATGCTGCTATATACCATATAG 882
OY 1629 atggattgagatgttgctctctcttaaggatctctctctttagaagaagatgacaacaagac 1688
Db 883 ATGATTTGGATGTTTGGCTCTCTTAGGATCTCTCTTAGAAGATGACACAAAGAC 942
OY 1689 agcttcaagaagacagatgagaatgcccacttgaggttcttctctgttctgtctgt 1748
Db 943 AGCTTCAAGAAAGACAGATGAGATGCCACTTGGGGCTTCTCTCTGCTGTGTGT 1002
OY 1749 gtgtgtgttctcaaaagtttcaagcttctctgttcttcttcttcttcttctatcatgt 1808
Db 1003 GTTGTGTGTCACAAAGTTTCAGCTTCTGTTCTTCTCTCTCTCTCTTATCATGT 1062
OY 1809 gtctcttcaaccttccaatataatgtttacaacatttctgtctgtctttaaactg 1868
Db 1063 GTCTCTTCACCTTTCATTAATATGTTACAAACATTTGCTGTAGTTAAACATG 1122
OY 1869 taaatgttgatgactcttcg 1889
Db 1123 TAAATGTTGATGATCTTTC 1143

RESULT 5
AP003214 138711 bp DNA PLN 13-APR-2001
LOCUS Oryza sativa genomic DNA, chromosome 1, BAC clone: OSJNBa0083M16,
DEFINITION complete sequence.
ACCESSION AP003214
VERSION AP003214.2 GI:13620983
KEYWORDS HTG.
SOURCE Oryza sativa (cultivar: Nipponbare) DNA, clone: OSJNBa0083M16.
ORGANISM Oryza sativa

REFERENCE
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
JOURNAL clone: OSJNBa0083M16
PUBLISHED Only in Database (2001) In Press
2 (bases 1 to 138711)
AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE Direct Submission
JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@abr.affrc.go.jp, URL: http://rtp.dna.affrc.go.jp/,

Tel: 81-298-38-7441, Fax: 81-298-38-7468
On Apr 12, 2001 this sequence version replaced gi:13027244.
The orientation of the sequence is from -21M13 to M13rev of the BAC
clone.

FEATURES
source Location/Qualifiers
1..138711
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="OSJNBa0083M16"

BASE COUNT 37474 a 31378 c 32147 g 37712 t
ORIGIN

Query Match 14.3%; Score 269.2; DB 8; Length 138711;
Best Local Similarity 57.7%; Pred. No. 2.9e-47;
Matches 548; Conservative 0; Mismatches 383; Indels 19; Gaps 3;

OY 727 ttgtctgagatgcaattccttaataaagcgtatgctgttcaaatgttgtgcaatgta 786
Db 90614 TTGATGACGATTCTATACCTACATCGAAGCTATGAAGAAGAAATTAATGTTGATGCA 90673
OY 787 aggcatacgaatgacactctcttcccaactgctccgagaagtatga-tgaacgtggt 846
Db 90674 AGGCTATGCTCTCTATTTGTGCCCTTCCACCTTATCTAGTACATGCTTGGAATGAT 90733
OY 847 ggaaccaaatgtttgttagacagaaatcgcgtgtaattcgtatc-ttgttacaatg 906
Db 90734 GGACACAGCTTATGTTATATACAGTAAGAGTGTGGTGGCCAAATGTAAC TGGTTTATCTGG 90793
OY 907 ccatcattcgttttctgtgtgtgtgtgtatattatgtatgtacagagtgcttcatgaca 966
Db 90794 CTTTATCTTATCTTGTGTTGAGTTGGAATTTACTGATGACAGAGTGTGATGACAGA 90853
OY 967 ttaagcctctcatalagatctccatgcccacatcaatctcacaagaagaaataacac 1026
Db 90854 TAAACCAATGTGTACAGATACGACACATACCATATTTTACAAACA GGAAGAAATACC 90913
OY 1027 tctctcaattgcggaagtgt-----ttcaagttgtcttctttagtctc 1074
Db 90914 TATCACCAATTTGACAGTAAATTTTATGAGCCGCTTAAATATGCAATTCATTTTAA 90973
OY 1075 tctaaagaattgtagcattgtttcttaccagaagaacttctgca-icagctgtgt 1134
Db 90974 TGTTCGCTGTGCAATGATATTAATGAGAACAGATTCATATTCGATVGCCTTATTTT 91033
OY 1135 actccaatccaatcttgcattcctatccataaaglaacagaaag-ctagaattalat 1194
Db 91034 ACCATGCTTCCACGCTTTTATGATATATCTGTGATTAATATATATTTCCCTGCTAA 91093
OY 1195 aaatgcaagctgcatctactcacaatalgtcagaagaactctga----ttaacaagagt 1250
Db 91094 TTAATGTGCTTAAACACTAGGATACCTTATTTAAGGCTCTTAAACCTTTTACCCGTAGG 91153
OY 1251 tttagatcttgttcttcttcttctgtctcgcgactgattggaatgacg-agaagttctt 1310
Db 91154 CCATACCAATTAATAAAAAAATGCAATGCGCTCCCAATGTATCAATTTTAA--CAATA 91210
OY 1311 atctactccctcggaggtgcatcttctgttaatccaagaagtacatctataatctgt 1370
Db 91211 CAATTTTCTTATTAAGTAACTCCCTTTGCTGTGAGAAATGTGATTTCTTCACCAAGCTTCT 91270
OY 1371 aacttcccttaagcttcttcttcaagggttgcatttccaccagtagacjgtaactcag 1430
Db 91271 GATTTATGCTCTTGCATGCTTCCAGACTAGCATTCATCCATGGAATGGGATTTTGGCAA 91330
OY 1431 gctgtaacgcatgtagatgctgtttatagtcgaatcattcatatcaatcatataggt 1490
Db 91331 GGCATACCGCATGTGTTGCGCTTACCTTATCCCAACACACTTCAGGACACACATGTGT 91390
OY 1491 ctltgttcatggaagcagatagacgcgaacatccatgactgtatc-actgcaaacatc 1550
Db 1550 ctltgttcatggaagcagatagacgcgaacatccatgactgtatc-actgcaaacatc 1550

Dd	91391	CTCTGTGTTATAGAGGCCCGGTGGACAACTAACATCCATGACGTCAATTCACGGCAAGSTT	91450
Oy	1551	tgggcagtaatgaggtgcagagatgccatagatccacacagatataaagatacaatcat	1610
Dd	91451	tgcccgcgatgagtgctgctgcgctatccacacattccattccatccatccgctccacactat	91510
Oy	1611	ggatcatataaccatcagatgagatctgagatgttggctcctctccttaggagatcc	1660
Dd	91511	ggccacattacacccgttgatggatggactcgatgtttggggcacccttgtagagagcc	91560

RESULT	6				
AF081794	AF081794	1155 bp	mRNA	PLN	19-MAY-1999
LOCUS	Nicotiana tabacum sterol-C(5)-desaturase mRNA, complete cds.				
DEFINITION	AF081794				
ACCESSION	AF081794.1	GI:4140397			
VERSION					
KEYWORDS					
SOURCE	Common tobacco.				
ORGANISM	Nicotiana tabacum				

REFERENCE	AUTHORS	TITLE	JOURNAL
1 (bases 1 to 1155)	Husselstein, T., Schaller, H., Gachotte, D. and Benveniste, P.	Delta ⁵ -sterol-C5-desaturase: molecular characterization and functional expression of wild-type and mutant alleles	Plant Mol. Biol. 39 (5), 891-906 (1999)

MEDLINE 99273992
REFERENCE 2 (bases 1 to 1155)
AUTHORS Benveniste, P.
TITLE Direct Submission
JOURNAL Submitted (30-JUL-1998) Plant Molecular Biology Institute, Centre

National de la Recherche Scientifique, 28 rue Goethe, Strasbourg 67084, France	Location/Qualifiers
1. .1155	source

```

/organism="Nicotiana tabacum"
/cultivar="xanthi SH6"
/db_xref="taxon:4097"
/tissue_type="calli derived from leaf protoplasts"
103. .918

```

```

/codon_start=1
/product="sterol-C5(6)-desaturase"
/protein_id="AAD04034.1"
/db_xref="GI:4140398"
/translation="MEDYKQFVEETSFYNRVLVGFMEPSWGRPLRPHLQGLRNR
GCVLLFTSGFLMCFYIHLKRNRYIPKATISNRKMLQISYAKMRYCALPSI
EMLINGNTCFARISDVGWLSVYIALVLVEGIIYMHNEHLMDIKPLVYLRH
HIIYIKNLSFAGLAFRLDGLIQAIVHVALLLVPMHFSHIALIFLEALMTA
HDCIHGKQFVPMGAGYHTIHNRTYIRNYSCHYITMDMMEGTLDLPVEEDAKM"

```

Query Match	9.9%	Score 186.2;	DB 8;	Length 1155;
Best Local Similarity	72.2%;	Pred. No. 1.5e-29;		
Matches 242;	Conservative 0;	Mismatches 93;	Indels 0;	Gaps 0

Oy 720 gtcgtctttgcctcagatcgcaatttccctaacaataaaggctatgctgtttgcaaatgttcgtg 779
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 gtatatatttccaaagatgccatcccctttciaaacaaagcaattgctcttgccaattatcagta 363

Oy 780 gcaatgaagcgtacgcactgtagcactctcttcaccactgctcccgagagtatgatgaa 839
 || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 364 gccattgaaggcatgactgtagtctgccccttcacatcacattttctgagtagaatattgaa 423

Qy	840	cgltgttgagccaatgcttctgacatagacaagaatctgcgcgattccgtatattgct	899
Db	424	AATGCATGCACAAATGTTTTCGACAACATAAGCGATGTTGCATGCGCTTTCCATATTATC	483

900 tacatcgcacatcatctcttcttcgttgcgttgcattatctgcagtcacagagagcct 959

Db	484	TATGGCGCTATTATTCTTTGTAAATAGTACGATTTGGGATCTACTGTGGATGCAATATGGAGTTG	543
Qy	960	catgacatcaagcctcctcctataaglatctccatgycacccatcatctacacacaagag	1019
Db	544	CATGACATTAATAACCTCTGTACAAATATCTCCATCTCTACACATCATATTATTAACAACAGCAA	603
Qy	1020	atacactctctccattgtgcccgtgaagtgttttca	1054
Db	604	AATACACTTTTCCCATTTGCTGTGGATTGGCAATTCGA	638

RESULT	7
AF099969	
LOCUS	AF099969 1148 bp mRNA
DEFINITION	Nicotiana tabacum sterol-C5(6)-desaturase homolog mRNA, complete cds.
ACCESSION	AF099969
VERSION	AF099969.1 GI:4426626
KEYWORDS	.
SOURCE	common tobacco.

REFERENCE
Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
1 (bases 1 to 114b)
Husselstein, T., Schaller, H., Gachotte, D. and Benveniste, P.
Delta1-sterol-5-desaturase: molecular characterization and
functional expression of wild-type and mutant alleles

JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE

1. Unclonal expression of wild-type and mutant alleles
Plant Mol. Biol. 39 (5), 891-906 (1999)
99273992
2 (bases 1 to 1148)
Benveniste, P.
Direct Submission

JOURNAL	Submitted (20-OCT-1996) Plant Molecular Biology Institute, C.N.R.S., 28 rue Goethe, Strasbourg 67083, France
FEATURES	Location/Qualifiers
SOURCE	1. 1148

```

/organism="Nicotiana glauca"
/cultivar="xanthi SH6"
/db_xref="taxon:4097"
/cell_type="protoplast (call)"
/tissue_type="leaf"

```

```

COS
127. .942
/function="introduces a double bond at C5 of
delta7-avenasterol resulting in
stigmastera 5,7,224(28)-triene-3beta-O1"
/codon_start=1
/product="sterol-C5(6)-desaturase homolog"
/protein_id="AAD20458.1"
/db_xref="GI:4426627"
/translation="MDDYLTLEETSFVYRVVLTGFLPDSWNGPLPHHFOGWLRL
GGLVLFISGFLMCFIYRLKRVVYIPKDIIPNRMLLIGVAMKAMPEYCALP
EYVINGWTFCSRISDVGWLSYLIYVAVLYEYVIMHEDLHDIKILYLYL

```

BASE COUNT	302 a	236 c	229 g	381 t
ORIGIN	HHIYKQNTLSPAGLAFHPLDGILQAVPHVALFLPEFTTHALLPLEATW7 HDCNAKAVPMGAGYHTIHTTYYRNNYGYTMMDMGCTLADPVEDEYKK"			

Query Match	9.4%	Score 178.2;	DB 8;	Length 1148;
Best Local Similarity	70.7%	Pred. No. 7.3e-28;		
Matches 237;	Conservative	0;	Mismatches 98;	Indels 0;
			Gaps	

[illegible]

```
Oy      780 gcaatgaagcgtatgcacatggttacactcttcctccaacctctcccgagaatatgatggaa   839  
         ||||| ||||||| | | ||||| | | ||||| ||||| |  
Db     388 GCCATGAAGCGTATGCCATTCTTTGTCCTCCGCACACTTTCTGGAGTAATGATGATTGTA    447
```

QY 840 cgtgcttggaaccaatcttctgcatagacgaattcgcctggaattctgtatttcgtc 899
||| ||||||| ||||| ||| ||| ||| ||| ||| |||

Db 448 AATGGGTGACAAAATGTTTCTCAAGAAATTAGCAGCTTGAGTGGCTTCTACCTTAT 507
 Oy 900 tacatcgcatctatctgttttgcgttgatgttattatgtatgacagaagctt 959
 Db 508 TATATGGCAGTTTATTTGGTAATAGATATTCGGCATTTACTGGATGATGCGAGTTG 567
 Oy 960 catgacataagcctctatlaagatctcctacatgacacccatcatatcacaagaagcag 1019
 Db 568 CATGACATAAAACCTCTGTACAAATATCTCCATGCTACACATCATATTTACAAACAA 627
 Oy 1020 aatacactctcacttgcgcgttaagtgcttca 1054
 Db 628 AATACACTCTCCCATTTGCTGGTTGGCTTCCA 662

RESULT 8
 AX155063 1381 bp DNA PAT 22-JUN-2001
 LOCUS AX155063
 DEFINITION Sequence 109 from Patent WO0138484.
 ACCESSION AX155063
 VERSION AX155063.1 GI:14536617
 KEYWORDS
 SOURCE Physcomitrella patens.
 ORGANISM Physcomitrella patens
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
 Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
 1 (bases 1 to 1381)
 Ierch, J., Renz, A., Ehrhardt, T., Reindl, A., Citpus, P., Bischoff, F.,
 Frank, M., Freund, A., Duenweg, E., Schmidt, R. M. and Reski, R.
 Moss genes from Physcomitrella patens encoding proteins involved in
 the synthesis of polyunsaturated fatty acids and lipids
 Patent: WO 0138484-A 109 31-MAY-2001;
 BASF Plant Science GmbH (DE)

FEATURES
 source Location/Qualifiers
 1..1381
 /db_xref="taxon:3218"
 40..1041
 /note="p004064012R"
 /codon_start=1
 /protein_id="CAC42667.1"
 /db_xref="GI:14536618"
 /translation="MASRGAVNMCALAVLWVAMSLSCMSADVEVYNVAFSSVVG
 GAKTSGVAVPANGSPYALFVEETRYNDVLGPMPLSSVRDIPTLTQMLNRY
 AGMLYFVSGIWLTYYSKMGHFPAGDIIPAKETIMLOITMMAKADPVYTGIDPIS
 EYMIEBGTWCFARIEDYGLTYGVYATYAVYFEGTYMMRELDHDKPLKHLHAT
 HHTYKONTLSPEFGLAHPIIDGLQCPHYIALFLPMHFTEHVLLECGVWTTNI
 HDCIDGVNWMGIMGAFHTIHTTYRHNGHYVMDLFTLRDPYERKATAVHKS"

BASE COUNT 323 a 331 c 351 g 376 t

Query Match 8.7%; Score 164.2; DB 6; Length 1381;
 Best Local Similarity 68.7%; Pred. No. 6.8e-25;
 Matches 226; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

Oy 726 ttgtcttgcgtatgcattccctacaaataagcgtatgcgtttgcaaatgtttgcaatg 785
 Db 421 TTCCCTGAGGTGACATACCCGCGAAGAGCCCAATATGCTCAATATCGGGTATAG 480
 Oy 786 aagcgtatgcataagcattcttccaaatgctccgaagatgataaagcgtatg 845
 Db 481 AAGGCTATGCGCATATACAGAGACTTCCCACTCTGTCGCAATATATGATTGAGCGGGG 540
 Oy 846 tggaccaaagtcttgcagacataagcgaatcgctgagatctgatttttacaac 905
 Db 541 TGGACCAAGTGTTCGCGCATCGAGATGTTGGTGGCTCAGCTATGAGCCTAGTC 600
 Oy 906 gcatctatctgttttgcgttgaagtgatattatgataagagaagattatgac 965
 Db 601 ATCCGCTACTTGCGAGTGTGAGATTGTGATCTATGGATGACAGAGAGCTTACGAT 660
 Oy 966 ataagcctctatlaagatctcactcaccacacatcatatcacaagaagataata 1025

Db 661 ATTAAGCCTTTGTACAAACATCTGCATGTCACCCACACACTATACATAAGCAAAACAGC 720
 Oy 1026 cctctccatctgcgcgttaagtgcttca 1054
 Db 721 CTATACACCGTTTGACAGTTTGGCGTTCCA 749

RESULT 9
 G71612/c 310 bp DNA STS 08-JUN-2001
 LOCUS G71612
 DEFINITION A62988234FM017 maize leaf DNA Zea mays STS genomic, sequence tagged
 site.
 ACCESSION G71612
 VERSION G71612.1 GI:14333297
 KEYWORDS STS.
 SOURCE Zea mays.
 ORGANISM Zea mays
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 310)
 Yang, Y.-J., Guo, L., Ashlock, D.A., Wen, T.J. and Schnable, P.S.
 3' UTR sequences of maize genes
 Unpublished

REFERENCE
 AUTHORS Schnable, P.S.
 TITLE Schnable Laboratory
 JOURNAL Iowa State University
 405 Agronomy Hall, Ames, IA 50011, USA
 Tel: 515-294-0975
 Fax: 515-294-2299
 Email: schnable@iastate.edu
 Primer A: CTAAATGACACCAAGCC
 Primer B: GATGTAATTCGCAAGCG
 PCR profile:
 Denaturation: 94 degrees C for 30 seconds
 Annealing: 60 degrees C for 45 seconds
 Polymerization: 72 degrees C for 90 seconds
 PCR cycles: 31
 Thermal cycler: Perkin Elmer TC
 Protocol:
 Template: 10-20 ng
 Primer: each 0.5 uM
 dNTPs: each 200 uM
 Taq Polymerase: 0.05 units/uL
 Total vol: 20 uL

Buffer:
 MgCl2: 2 mM
 KCl: 50 mM
 Tris-HCl: 20 mM
 pH: 8.4.

FEATURES
 source Location/Qualifiers
 1..310
 /organism="Zea mays"
 /strain="DE811"
 /db_xref="taxon:4577"
 /clone_lib="maize leaf DNA"
 /note="PCR products amplified from genomic DNA"

BASE COUNT 80 a 81 c 87 g 62 t

Query Match 7.0%; Score 132.2; DB 11; Length 310;
 Best Local Similarity 73.0%; Pred. No. 4.5e-18;
 Matches 170; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Oy 1428 cagcgtatcgcgcatatgtagcgtgttattatagtgcaattcatc:acaactcatata 1487
 Db 309 CAAGCAATGCCCATGTGCTTGGCTCTCTCTCCCAACGACATT:AGGACGCACATC 250

LOCUS	CNS0180K	840 bp	MRNA	PLN	02-SEP-1999
DEFINITION	Botrytis cinerea strain T4 cDNA library under conditions of nitrogen deprivation.				
ACCESSION	AL106675				
VERSION	AL106675.1 GI:5824962				
KEYWORDS	cDNA library; nitrogen deprivation.				
SOURCE	Botryotinia fuckeliana.				
ORGANISM	Botryotinia fuckeliana Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Sclerotiales; Botryotinia.				
REFERENCE	1 (bases 1 to 840)				
AUTHORS	Biton,F., Lewis,C., Fortini,D., Pradier,J.M. and Brygoo,Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr, 78026 Versailles, France				
REFERENCE	2 (bases 1 to 840)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (01-SEP-1999) Genoscope - Centre National de Sequencage : CP 5706 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)				
COMMENT	- Web : www.genoscope.cns.fr The cDNA library to be analyzed within the framework of this project was created using a Botrytis cinerea strain which was grown under conditions of nitrogen deprivation, which is the normal situation for B. cinerea during its development on its host plant. The library was produced in an oriented direction, in the pSII vector.				
FEATURES	Location/Qualifiers				
source	1..840 /organism="Botryotinia fuckeliana" /strain="T4" /db_xref="taxon:40559" /note="Genoscope sequence ID : W33E091"				
BASE COUNT	151 a 23 c 79 g 539 t 48 others				
ORIGIN					
Query Match	3.0%; Score 57.6; DB 8; Length 840;				
Best Local Similarity	46.3%; Pred. NO. 0.03;				
Matches 186; Conservative 0; Mismatches 216; Indels 0; Gaps 0;					
QY	377	tttcacatttgatcactatgtcctaactgcctctcatgtatcgattttcaattaa	436		
DB	249	tt	308		
QY	437	ggaagaagggttcctcgctgactgacaglaattggattgagtgtgataagttcatgt	496		
DB	309	tt	368		
QY	497	tgcattatgattatatttgatcatattccactcaagaagattgaacagttagtgacta	556		
DB	369	tt	428		
QY	557	tataagatttggcgaacacatgagaagtcgtacacatttgaagttgaatttctactgt	616		
DB	429	tt	488		
QY	617	ccatttaagtcacacttaaatgttggtgagaagtgatgtctactcttcagacacatctt	676		
DB	489	tatt	548		
QY	677	ttctgctctcgcgaacatcgtctcattgtaacatttttggtctgttttgcttcaga	736		
DB	549	tt	608		
QY	737	tgaacttcacacaataaagcctacgttggttgcaaatgttgtt	778		
DB	609	tt	650		

LOCUS	GENBANK	696 bp	mRNA	PLN	02-SEP-1999
DEFINITION	Botrytis cinerea strain T4 cDNA library under	conditions of			
ACCESSION	AL114081				
VERSION	AL114081.1	GI:5828700			
KEYWORDS	cDNA library; nitrogen deprivation.				
SOURCE	Botryotinia fuckelliana.				
ORGANISM	Botryotinia fuckelliana				
REFERENCE	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Leotiomycetes; Helotiales; Sclerotiales; Botryotinia.				
AUTHORS	1 (bases 1 to 696)				
TITLE	Biton, F., Lewis, C., Fortini, D., Pradier, J.M. and Brygoo, Y.				
JOURNAL	Direct Submission				
REFERENCE	Submitted (01-SEP-1999) Phytopathologie, INRA, route de St Cyr,				
AUTHORS	78026 Versailles, France				
TITLE	2 (bases 1 to 696)				
JOURNAL	Genoscope.				
COMMENT	Direct Submission				
	Submitted (01-SEP-1999) Genoscope - Centre National de Sequences				
	CP 5706 91057 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)				
	- Web : www.genoscope.cns.fr				
	The cDNA library to be analyzed within the framework of this				
	project was created using a Botrytis cinerea strain which was grown				
	under conditions of nitrogen deprivation, which is the normal				
	situation for B. cinerea during its development on its host plant.				
	The library was produced in an oriented direction, in the pBSII				
	vector.				
FEATURES	Location/Qualifiers				
source	1..696				
	/organism="Botryotinia fuckelliana"				
	/strain="T4"				
	/db_xref="taxon:40559"				
	/note="Genoscope sequence ID : W62C01.1"				
BASE COUNT	32 a 13 c 124 g 513 t 14 others				
ORIGIN					
Query Match	3.0%; Score 57.4; DB 8; Length 696;				
Best Local Similarity	47.5%; Pred. No. 0.034;				
Matches	169; Conservative 0; Mismatches 187; Indels 0; Gaps 0;				
OY	377 ttccactcttgatcactactatgcttaacgcctctctatgtagtgcgaatttcaattaa	436			
Db	334 ttt	393			
OY	437 ggaagaagggtctctcgcgtactgcaacgaatttgagttgagtgatgaatcagtt	496			
Db	334 ttt	453			
OY	497 tgcattatgatattatgtagcatattctccatcacaaggatgaacugtagtgcta	556			
Db	454 ttt	513			
OY	557 tataagttttgtagcaacacatgagaagtcgtacatcttgaagttgnaatttctactg	616			
Db	514 ttt	573			
OY	617 ccatttaagtcacactaaatgtttgttggaagtgtgtctacttcagacacatcttt	676			
Db	574 ttt	633			
OY	677 ttctgctctctcagacacgtcgtactgaattggaacatttttgctcttttgcctt	732			
Db	634 ttt	689			
RESULT	14				
LOCUS	AB016248	1721 bp	mRNA	ROD	18-FEB-2000
DEFINITION	Mus musculus mRNA for sterol C-5-desaturase, complete cds.				
ACCESSION	AB016248				
VERSION	AB016248.1	GI:3721883			
KEYWORDS	sterol-C5-desaturase; C5D.				

SOURCE	Mus musculus (strain:B6B/C) adult male cDNA to mRNA.
ORGANISM	Mus musculus
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi: Mammalia: Eutheria: Rodentia: Scuriognathi: Muridae: Murinae: Mus.
AUTHORS	1 (sites)
TITLE	Nishi,S., Nishino,H. and Ishibashi,T.
JOURNAL	cDNA Cloning Of the mammalian sterol C5-desaturase and the
MEDLINE	expression in yeast mutant
REFERENCE	Biochim. Biophys. Acta 1450 (1-2), 106-108 (2000)
AUTHORS	20246294
TITLE	2 (bases 1 to 1721)
JOURNAL	Nishino,H.
MEDLINE	Direct Submission
REFERENCE	Submitted (09-JUL-1998) to the DDBJ/EMBL/Genbank databases. Hideoaki
AUTHORS	Nishino, Hokkaido University School of Medicine, Department of
TITLE	Biochemistry, N15W Kita-Ku, Sapporo, Hokkaido 060-8638, Japan
JOURNAL	(E-mail:hideoaki.nishino@med.hokudai.ac.jp, Tel:+81-11-706-5047,
MEDLINE	Fax:+81-11-706-5169)

FEATURES	SOURCE	LOCATION/QUALIFIERS
		1..1721
		/organism="Mus musculus"
		/strain="Balb/c"
		/db_xref="taxon:10090"
		/dev_stage="adult"
		/sex="male"
		171..1070
gene		/gene="C5D"
		171..1070
CDS		/gene="C5D"
		/EC_number="1.3.3.2"
		/codon_start=1
		/product="sterol-C5-desaturase"
		/protein_id="BAA33730.1"
		/db_xref="GI:3721884"
		/translation="MDLVLSAADYFFTPVYVPATWEDNIIRQTISLLIVTNGAYI
		LYFPCALSYFYVYDHSILMKHPDLKNQVSREIYFVKSLEWISIPVLSLELRQY
		SLKVDIGDPENGWIIHLWVSFLKFTDMLIVIRHGLHRLVYKRIRKPHHIMKIP
		TPFSAHFHPVDGFLQSLPYHIYFVPEPLKHVVYLGIVLVNWTISIHODEPVPDI
		LRPINSAAHTDHHMFEDYNGQYFTLMDRIGGSFRKPPSEFGKGFHSYVKNTEKE
		SNSPRAEGCKKAKYSNDEFTKNK
BASE COUNT	477 a	399 c 376 g 469 t
ORIGIN		

Query Match	3.0%;	Score 56.6;	DB 10;	Length 1721;
Best Local Similarity	50.7%;	Pred. No. 0.048;		
Matches 171;	Conservative	0;	Mismatches 154;	Indels 12;
				Gaps 1;
QY	1388	gtctcaagagccttgatcttcacccagtagagaggaactcaagctgtgacgcgcatgcat	1447	
Db	656	gntttcagatgcattgcttttttccacctgtgagacgcgtctcttcagagctgcccattacat	715	
QY	1448	agcgctgtttatagtgcgaatcttcacatctacatatataggtctttgtttcabygaagc	1507	
Db	716	ataccctctttgtctttccacatgcacacaaagctgctacacttaggtttatagctttggttaa	775	
QY	1508	gatacgcgcgcggaaacatccatcagactgcatacabygaacaatctgcgcgaatg-----	1562	
Db	776	tgtcttgacacattttttatttattcattgattgatttttggcttccccagantctttaagccatt	835	
QY	1563	-----ggtcgaggaataccatagatatacaccacacagacatacaagcaataactatgltca	1615	
Db	836	tatttaagggggtcagctcacacacacagaccacacacattgtttctttgactttaactatgagaca	895	
QY	1616	ttataccalatgatalgataltgagatgttgcgctctcttagagatccctctttagaagaaga	1675	
Db	896	gtattttcacatttgttggtatagattgagagcctcttttaaacatctcttccctcttttgaaag	955	
QY	1676	tgacacaacaagacagcttccaagaaagcagagtgagaa	1712	
Db	956	gaaaagaccacattagttacgtgaamaacatgacacaa	992	

RESULT	15
CNS05TC1/c	
LOCUS	CNS05TC1 127709 bp DNA HTG 25-MAY-2000
DEFINITION	Homo sapiens chromosome 14 clone C-2216H20, *** SEQUENCING IN
PROGRAMS	***, 2 ordered pieces,
ACCESSION	AL352983
VERSION	AL352983.2 GI:9213046
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE	human;
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 127709)
TITLE	Genoscope.
JOURNAL	Direct Submission
COMMENT	Submitted (23-MAY-2000) to the EMBL/GenBank/DBJ databases on Jul 15, 2000 this sequence replaced gi:8218012.

represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continue. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage, etc. . . even if efforts are made to eliminate these contaminating sequences. The following BAC sequence is oriented from the T7 to the SP6 end.

Contigs composition :
9499 bp contig from 1 to 9499
118110 bp contig from 9600 to 127709

Overall quality chart :		
Range	:	bases
0	:	312
1	-	481
10	-	505
20	-	1404
30	-	4578
40	-	5806
50	-	7495
60	-	17934
70	-	37613
80	-	33579
90	-	17982

```

Percentage of bases with a quality value >= 40 : 94 %.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 9500 9599: gap of 100 bp
* 9600 .127709: contig of 118110 bp in length.

```

FEATURES	SOURCE	location/Qualifiers
		1. 127709
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="14"
		/clone="C-2216H20"
		/clone_11b="calTech-D"
BASE COUNT	36584 a 27636 c 26616 g 36396 t	277 others
ORIGIN		
Query Match	3.0%;	Score 56.6; DB 2;
Best Local Similarity	34.3%;	Length 127709;
Matches 158; Conservative	0; Mismatches 303; Indels 0; Gaps 0;	

```

Db 10196 TTCCCCGATCGTATTGAAATACNCCCCCGCTGNNNNNNNNNNNNNNNNNNNN 10137
QY 337 ttaccttaaatcaacgcttaccctcccaagctcgcacttcttcttaccacta 396
Db 10136 NNNNNNNNTNNNNNNNTTTTNTNNNTTNTNNNTTNTNNNTTNTNNNTTNT 10077
QY 397 ttgcttaacgccttctatgttaccatcttcaatlaagaagaaggcttcctcgtc 456
Db 10076 TTTTGTGTTTTTTTGTGTTTTTTTTTTTTTTTTTTTTNTNTNTNTTTTTTTNNN 10017
QY 457 actglacaglaattlgactlgactlgactlgactlgactlgactlgactlgact 516
Db 10016 NTTTTNTTNTTNTTNTTNTTNNNNNNNNNTNTNNNNNTTNTNTNTNTNTT 9957
QY 517 gcatatctcactcaaggaattgaacagtlagtgctatataaagtttctgcaacca 576
Db 9956 TNNTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNN 9897
QY 577 atgaagaagtcgtacacacttgaagttgaattctacttgcacatlaagtcacact 636
Db 9896 TTTTNTNTTTTTTTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 9837
QY 637 tgttctgaagtgacttgcacttgcagacacacttcttctcctcctgagactct 696
Db 9836 NTTTTTTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 9777
QY 697 gtctttagttgaacacttcttgcgtcttgccttcagat 737
Db 9776 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 9736

```

Search completed: April 5, 2002, 05:32:43
 Job time: 25017 sec

THIS PAGE BLANK (USPTO)